The University of San Francisco

USF Scholarship: a digital repository @ Gleeson Library | Geschke Center

Chemistry Faculty Publications

Chemistry

2000

Characterization of Chimeric Lipopolysaccharides from Escherichia coli Strain JM109 Transformed with Lipooligosaccharide Synthesis Genes (lsg) from Haemophilus influenzae

Nancy J. Phillips

T. J. Miller

J. J. Engstrom

William Melaugh
University of San Francisco, melaugh@usfca.edu

R. McLaughlin

See next page for additional authors

Follow this and additional works at: http://repository.usfca.edu/chem_fac

Part of the Biology Commons, and the Chemistry Commons

Recommended Citation

Authors

Nancy J. Phillips, T. J. Miller, J. J. Engstrom, William Melaugh, R. McLaughlin, M A. Apicella, and B W. Gibson

Characterization of Chimeric Lipopolysaccharides from *Escherichia coli* Strain JM109 Transformed with Lipooligosaccharide Synthesis Genes (*lsg*) from *Haemophilus influenzae**

Nancy J. Phillips‡, Theresa J. Miller§, Jeffrey J. Engstrom‡, William Melaugh‡, Robert McLaughlin¶, Michael A. Apicella§, and Bradford W. Gibson‡‖

From the ‡Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446, the ¶Department of Microbiology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, and the §Department of Microbiology, University of Iowa, Iowa City, Iowa 52242

Previously, we reported the expression of chimeric lipopolysaccharides (LPS) in *Escherichia coli* strain JM109 (a K-12 strain) transformed with plasmids containing *Haemophilus influenzae* lipooligosaccharide synthesis genes (*lsg*) (Abu Kwaik, Y., McLaughlin, R. E., Apicella, M. A., and Spinola, S. M. (1991) *Mol. Microbiol.* 5, 2475–2480). In this current study, we have analyzed the *O*-deacylated LPS and free oligosaccharides from three transformants (designated pGEMLOS-4, pGEM-LOS-5, and pGEMLOS-7) by matrix-assisted laser desorption ionization, electrospray ionization, and tandem mass spectrometry techniques, along with composition and linkage analyses. These data show that the chimeric LPS consist of the complete *E. coli* LPS core structure glycosylated on the 7-position of the non-reducing terminal branch heptose with oligosaccharides from *H. influenzae*. In pGEMLOS-7, the disaccharide Gal1→3GlcNAc1→ is added, and in pGEMLOS-5, the structure is extended to Gal1→4GlcNAc1→3Gal1→3GlcNAc1→. PGEMLOS-5 LPS reacts positively with monoclonal antibody 3F11, an antibody that recognizes the terminal disaccharide of lacto-*N*-neotetraose. In pGEMLOS-4 LPS, the 3F11 epitope is apparently blocked by glycosylation on the 6-position of the terminal Gal with either Gal or GlcNAc. The biosynthesis of these chimeric LPS was found to be dependent on a functional *wecA* (formerly *rfe*) gene in *E. coli*. By using this carbohydrate expression system, we have been able to examine the functions of the *lsg* genes independent of the effects of other endogenous *Haemophilus* genes and expressed proteins.

Capsular strains of *Haemophilus influenzae* type b (Hib)[1] are

‖ To whom correspondence and reprint requests should be addressed: School of Pharmacy 926-S, 513 Parnassus Ave., University of California, San Francisco, CA 94143-0446. Tel.: 415-476-5320; Fax: 415-476-0688; E-mail: gibson@socrates.cgl.ucsf.edu.

[1] The abbreviations used are: Hib, *H. influenzae* type b; CID, collision-induced dissociation; ESI-MS, electrospray ionization-mass spectrometry; Hep, L-*glycero*-D-*manno*-heptose; Hex, hexose; HexNAc, *N*-acetylhexosamine; HPLC, high performance liquid chromatography; Kdo, 3-deoxy-D-*manno*-octulosonic acid; LOS, lipooligosaccharide(s); LPS, lipopolysaccharide(s); *lsg*, lipooligosaccharide synthesis genes;

responsible for various invasive and bacteremic infections in humans, including meningitis and pneumonia. The surface lipooligosaccharides (LOS) of Hib are known to be important factors in microbial virulence and pathogenesis (1–3). Structural studies of Hib LOS from wild-type (4, 5) and mutant strains (6–8) have shown that the LOS contains a conserved heptose trisaccharide core that can be extended with additional sugars on each heptose. However, efforts to correlate defined LOS structures with specific biological functions have been hindered by the high degree of microheterogeneity and antigenic variability of Hib LOS. Additionally, some Hib LOS epitopes phase vary at high frequencies due to slip-strand mispairing (9), adding a further level of structural complexity to wild-type Hib LOS populations.

To facilitate the elucidation of LOS biosynthesis and to determine the relationship of LOS structures to the biology and pathophysiology of *Haemophilus* infections, our laboratory has cloned a gene cluster from Hib strain A2 containing LOS synthesis genes (*lsg*) (10). The *lsg* loci are contained within a 7.4-kb DNA fragment, consisting of seven complete open reading frames (ORFs) (11). This region is one of several distinct loci found in the genome sequence of Hib strain Rd (12) that have been associated with lipopolysaccharide (LPS) biosynthesis (13). Recently, a series of isogenic mutants of Hib A2 was generated by transposon mutagenesis of the *lsg* region using minitransposon m-Tn*3*(Cm) (14). Several of the transposon mutants produced much simpler LOS mixtures than the wild-type Hib A2 LOS and no longer reacted with one or more of the monoclonal antibodies (mAbs) recognizing phase-varying or stable Hib epitopes (14). Structural characterization of the LOS from two transposon mutants with m-Tn*3*(Cm) insertions in *lsgD* and *lsgE* has suggested that mutation of these ORFs prevents the organism from synthesizing many of the higher molecular weight structures found in wild-type Hib A2 LOS (6). Strain 281.25, whose mutation was in *lsgD*, produced only the basal Hib A2 LOS structures containing glucose (Glc), heptose (Hep), and 3-deoxy-D-*manno*-octulosonic acid (Kdo). This strain did not express the Kdo-related 6E4 epitope characteristic of Hib LOS, or the 3F11 epitope, which consists of a terminal *N*-acetyllactosamine that is common to many strains of *Haemophilus* and neiserial LOS and has furthermore been impli-

LSIMS, liquid secondary ion mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization-time of flight mass spectrometry; (M − *n*H)^{*n*−}, deprotonated molecular ion; MS/MS, tandem mass spectrometry; ORF, open reading frame; PEA, phosphoethanolamine; r-LPS, rough LPS; PAGE, polyacrylamide gel electrophoresis; kb, kilobase pair; *O*-DPLA, *O*-deacylated diphosphorylated lipid A.
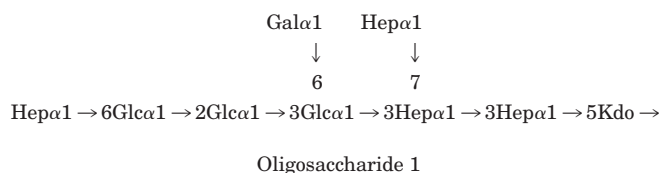
TABLE I
Bacterial strains and vectors

Strain/plasmid	Relevant characteristics	Ref./source
E. coli		
JM109	*recA, supE, hsR, Δ(lac-pro)*	48
21548	K-12, *thr-1,leuB6 Δ(gpt-proA06, hisG4, argE3, thi-1, lacY1, ara-14, galK2, xyl-5, mtl-1, mgl-51, rpsL31, kdgK51, supE44, wecA::Tn10–48)*	19
H. influenzae		
A2	Parental strain	10
Plasmids		
pGEM3Zf+	Ap^R	Promega Biotech
pGEMLOS-4	Ap^R, contains 7.4-kb *Bam*HI-*Pst*I DNA *H. influenzae lsg* locus	This study
pGEMLOS-5	Ap^R, contains 5.5-kb *Hin*dIII-*Pst*I DNA *H. influenzae lsg* locus	This study
pGEMLOS-7	Ap^R, contains 2.8-kb *Sph*I-*Pst*I DNA *H. influenzae lsg* locus	This study
pGEMG	Ap^R, contains ORF *lsgG*	This study
pGEMF	Ap^R, contains ORF *lsgF*	This study

cated in host mimicry (15). Strain 276.4, whose mutation was in *lsgE*, produced similar low molecular weight LOS components that failed to react with mAbs 6E4 and 3F11, and one high molecular weight species that terminated in sialyl-*N*-acetyllactosamine. When strain 276.4 was treated with neuraminidase to remove sialic acid from the sialylated LOS glycoform(s), the strain could now be recognized by mAb 3F11. Since both strains failed to produce any of the high molecular weight wild-type LOS structures containing galactose (Gal) or to express the P^k epitope (Galα1→4Galβ1→4Glc1→) associated with wild-type *H. influenzae* LOS (4, 15), these studies suggested that *lsgD* and *lsgE* are most probably involved in the transfer of galactose to the LOS.

In a related set of experiments, DNA from the *lsg* region of Hib strain A2 was used to construct a genomic library in the lambda bacteriophage EMBL3 (10). Twenty six phage clones were prepared that expressed Hib LOS oligosaccharide epitopes in *Escherichia coli* strain LE392. The phage transformant designated EMBLOS-1 produced a chimeric LPS with an ~1.4-kDa oligosaccharide added to the ~4.1-kDa LPS of *E. coli* LE392 (10). mAb 6E4, which recognizes a component in the Hib A2 LOS mixture, also recognized the novel 5.5-kDa component in the chimeric LPS. Restriction fragments of EMBLOS-1 were used to make a series of plasmids that modified *E. coli* JM109 to give clones that produced a proposed chimeric series of higher mass LPS species (16). The transformants pGEM-LOS-4, pGEMLOS-5, and pGEMLOS-7 generated modified or chimeric LPS of ~5.5, 5.1, and 4.5 kDa, respectively (16). All three apparently modified the ~4.1-kDa LPS species from *E. coli*, although only the LPS from pGEMLOS-4 expressed the 6E4 epitope. The LPS from strain pGEMLOS-5 was found to react positively with mAb 3F11, suggesting the presence of terminal *N*-acetyllactosamine.

The epitope recognized by mAb 6E4 is also present in the LOS of *H. influenzae* nontypable strain 2019 (17), as well as the LPS from *Salmonella minnesota* Re mutant. Binding of this monoclonal antibody to *H. influenzae* LOS can be inhibited by Kdo and the Kdo trisaccharide from the Re mutant of *S. minnesota* (10). Because the 6E4 epitope has been associated with the core of *Haemophilus* LOS, it was originally proposed that the chimeric structures expressed in *E. coli* might arise from the addition of a *Haemophilus* core structure to an acceptor residue of the *E. coli* ~4.1-kDa LPS species. Recently, a revised structure of the *E. coli* K-12 core region was reported that contains a heptose trisaccharide

inner core and a fourth heptose present on the terminus of the main oligosaccharide branch (18) as shown in Oligosaccharide 1.

$$\begin{array}{cc} \text{Gal}\alpha 1 & \text{Hep}\alpha 1 \\ \downarrow & \downarrow \\ 6 & 7 \end{array}$$

Hepα1 → 6Glcα1 → 2Glcα1 → 3Glcα1 → 3Hepα1 → 3Hepα1 → 5Kdo →

Oligosaccharide 1

In an effort to establish the chemical structures of the 0.4, 1.0, and 1.4-kDa modifications to the *E. coli* core and determine the *E. coli* acceptor residue, we report here the partial characterization of the LPS from *E. coli* strain JM109 and the three clones pGEMLOS-7, pGEMLOS-5, and pGEMLOS-4 using composition analysis, linkage analysis, and mass spectrometry.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Vectors—*H. influenzae* strain A2 was originally isolated from the spinal fluid of a child with meningitis. Hib A2 was grown on chocolate agar supplemented with IsoVitaleX or brain heart infusion agar supplemented with 4% Fildes reagent (Difco) at 35 °C in 5% CO$_2$ atmosphere. *E. coli* strains were routinely cultured at 37 °C using LB agar or broth with appropriate antibiotics (16). Vectors used in these studies were previously described (10, 16). *E. coli* K-12 strain 21548 which has a Tn10 insertion in *wecA* (formerly *rfe*) was obtained from Dr. C. Whitfield. This strain was originally described by Meier-Dieter and co-workers (19).

Transformation of E. coli Strains JM109 and 21548 with Plasmids Containing Restriction Fragments from the lsg Locus—*E. coli* strains JM109 and 21548 were transformed with the plasmid pGEM3Zf+ into which different DNA restriction fragments from the *H. influenzae* strain A2 *lsg* locus had been ligated (see Table I and Fig. 1) (20). The plasmid pGEMLOS-4 contained a 7.4-kb *Bam*HI-*Pst*I fragment of DNA that contained all seven open reading frames (A–G) comprising the *lsg* locus (11). The plasmid pGEMLOS-5 contained a 5.5 kb *Hin*dIII-*Pst*I fragment of DNA comprising five open reading frames (C–G) of the *lsg* locus. The plasmid pGEMLOS-7 contained a 2.8-kb *Sph*I-*Pst*I fragment of DNA comprising two open reading frames (F and G) of the *lsg* locus. In addition, both strains were transformed with pGEMG and pGEMF. The plasmid pGEM3zf+ without an insert was transformed into strain JM109. This strain and the LPS isolated from it were termed pGEM.

Isolation and SDS-PAGE Analysis of LPS—LPS from *E. coli* strain JM109, strain 21548, and the chimeric strains was prepared by the extraction procedure of Darveau and Hancock (21). The LPS were separated by SDS-PAGE in resolving gels containing 15% acrylamide

FIG. 1. **The *lsg* region of *H. influenzae* DNA.** *A,* diagram of the eight ORFs. *B,* locations of m-Tn*3*(Cm) insertion sites (6). *C,* restriction maps of the EMBLOS-1 subclones that modified the *E. coli* JM109 LPS (adapted from Ref. 16).
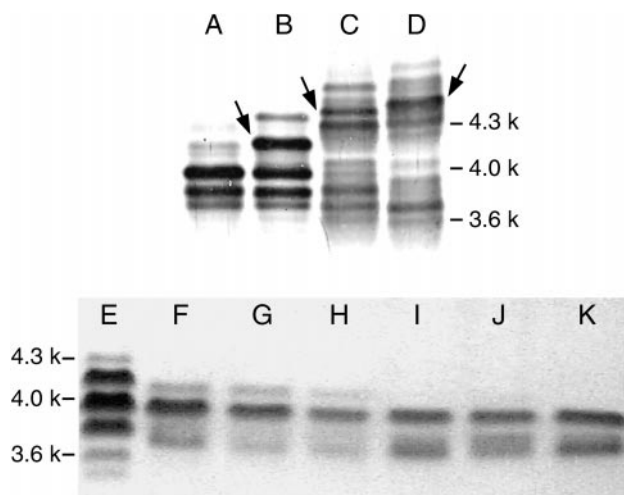
FIG. 2. *Top,* SDS-PAGE gel of the LPS from *E. coli* strain JM109 (pGEM) and the three chimeric strains. Lanes are as follows: *A,* pGEM; *B,* pGEMLOS-7; *C,* pGEMLOS-5; and *D,* pGEMLOS-4. Molecular weight markers indicated are the calculated masses of the LOS from *N. gonorrhoeae* strain PID2 (47). The apparent sizes of the most abundant moieties (labeled with *arrows*) added to the main *E. coli* LPS band at ~3.9 kDa, *i.e.* ~0.3, 0.5, and 0.6 kDa for pGEMLOS-7, -5, and -4, respectively, appear smaller than those originally suggested by SDS-PAGE (16) but closer to masses determined by mass spectrometry. *Bottom,* SDS-PAGE gel of the LPS produced when the *lsg* genes are transformed into *E. coli* strain 21548, a *wecA*− mutant. Lanes are as follows: *E,* PID2 from *N. gonorrhoeae; F,* pGEMrfe; *G,* pGEMGrfe; *H,* pGEMFrfe; *I,* pGEMLOS-7rfe; *J,* pGEMLOS-5rfe; *K,* pGEMLOS-4rfe.

fied LPS structures which added ~1.4, 1.0, and 0.4-kDa moieties, respectively, to the ~4.1-kDa *E. coli* core (16). In more recent SDS-PAGE gels, the modified LPS was further resolved and no longer appears as a dominant single band (Fig. 2, *top*). Rather, the LPS appear to consist of several bands whose estimated sizes differ slightly from those reported originally (see legend, Fig. 2). When tested against two mAbs that recognize *H. influenzae* epitopes, pGEMLOS-4 LPS reacted positively with mAb 6E4 (which recognizes a Kdo-related epitope in *H. influenzae*), and pGEMLOS-5 LPS reacted positively with mAb 3F11 (which recognizes the terminal disaccharide of lacto-*N*-neotetraose, Galβ1→4GlcNAcβ1→3Galβ1→4Glc1→). The LPS from pGEMLOS-7, which contains the smallest modifica-
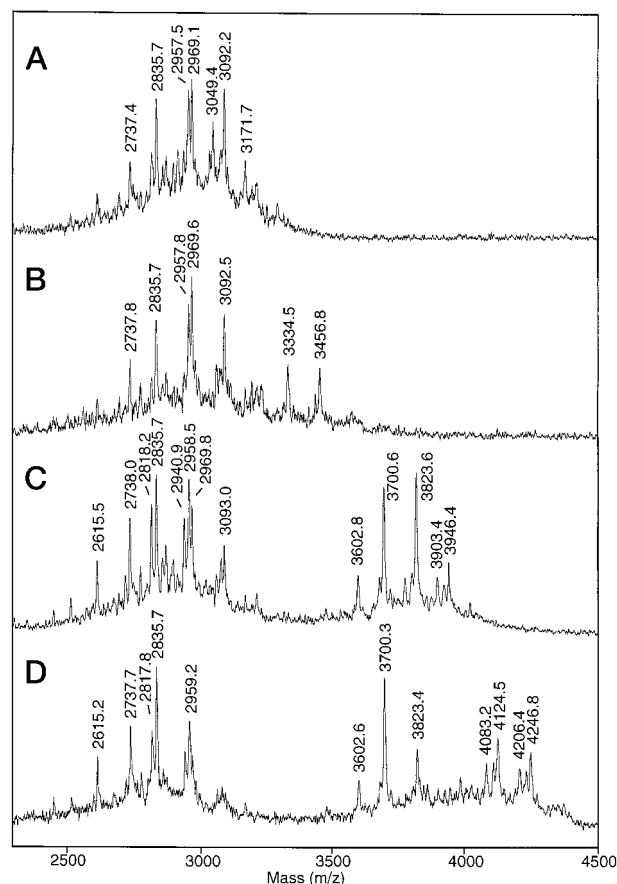
FIG. 3. **Negative ion MALDI-TOF spectra of the *O*-deacylated LPS from pGEM (*A*), pGEMLOS-7 (*B*), pGEMLOS-5 (*C*), and pGEMLOS-4 (*D*) recorded under delayed extraction conditions (30).**

tion, did not react with either antibody. Thus, the antibody recognition data and LPS size differences suggested that the *lsg* region contains a cluster of at least three genetic loci whose products sequentially modify *H. influenzae* LOS (16). To confirm the chimeric nature of the pGEMLOS-4, pGEMLOS-5, and pGEMLOS-7 LPS and identify *H. influenzae* epitopes being expressed, LPS from these clones and from the parent *E. coli* strain transformed with a plasmid lacking an insert (designated pGEM) were analyzed by various chemical and mass spectrometric methods as described below.

In order to define better the molecular mechanisms involved in the assembly of the chimeric LPS, we cloned the same plasmids as above plus plasmids containing only ORF *lsgG* and *lsgF* into strain 21548 which contains a Tn10 insertion in *wecA* (formerly *rfe*). The presence of a functional *wecA* is necessary for the synthesis of *O*-polysaccharides containing GlcNAc. The SDS-PAGE in Fig. 2 (*bottom*) demonstrates that a functional *wecA* is necessary for generation of the chimeric LPSs by the pGEMLOS-7, -5, and -4 plasmids.

Analysis of O-Deacylated LPS by MALDI-TOF—For initial screening of LPS molecular weights and heterogeneity by mass spectrometry, small aliquots of LPS from pGEM, pGEMLOS-4, pGEMLOS-5, and pGEMLOS-7 were treated with anhydrous hydrazine to remove *O*-linked fatty acids from the lipid A moiety. The negative ion MALDI-TOF spectra of these *O*-deacylated LPS samples (30) are shown in Fig. 3. Although all of the samples are clearly quite heterogeneous, the MALDI-TOF spectra easily show how progressively larger molecular weight LPS are made by the transformants as a function of *H. influenzae* ORFs present. The pGEM *O*-deacylated LPS sample

TABLE II

Molecular weights and proposed compositions of the O-deacylated LPS from pGEM, pGEMLOS-7, pGEMLOS-5, and pGEMLOS-4

Relative ion abundances are given in parentheses. P indicates phosphate.

pGEM	pGEMLOS-7	pGEMLOS-5	pGEMLOS-4	Calc. M_r	Proposed compositions						
					HexNAc	Hex	Hep	P	PEA	Kdo	O-DPLA
			Chimeric structures								
			4247.8 (40)	4247.6	3	6	4	2	1	3	1
			4207.4 (26)	4206.6	2	7	4	2	1	3	1
			4125.5 (46)	4124.6	3	6	4	2		3	1
			4084.2 (27)	4083.6	2	7	4	2		3	1
		3947.4 (38)		3947.3	2	6	4	2	2	2	1
		3904.4 (25)		3904.3	2	6	4	3	1	2	1
		3824.6 (100)	3824.4 (40)	3824.3	2	6	4	2	1	2	1
		3701.6 (90)	3701.3 (100)	3701.2	2	6	4	2		2	1
		3603.8 (31)	3603.6 (26)	3604.1	2	6	4	2	1	1	1
	3457.8 (41)			3458.9	1	5	4	2	1	2	1
	3335.5 (43)			3335.9	1	5	4	2		2	1
			Core structures								
3172.7 (43)				3173.6		4	4	3	1	2	1
3093.2 (97)	3093.5 (74)	3094.0 (42)		3093.6		4	4	2	1	2	1
3050.4 (67)				3050.6		4	4	3		2	1
2970.1 (100)	2970.6 (100)	2970.8 (68)		2970.6		4	4	2		2	1
2958.5 (90)	2958.8 (76)	2959.5 (87)	2960.2 (60)	2959.5	3	3	2	1		3	1
		2941.9 (56)		2942.3	3	3	3	2		2	1
2836.7 (87)	2836.7 (68)	2836.7 (90)	2836.7 (99)	2836.4	3	3	2			3	1
		2819.2 (70)	2818.8 (49)	2819.3	3	3	3	1		2	1
2738.4 (40)	2738.8 (41)	2739.0 (62)	2738.7 (54)	2739.3	3	3	2	1		2	1
		2616.5 (39)	2616.2 (39)	2616.2	3	3	2			2	1

contains several species in the range of 2738–3172 Da,[2] representing the major *E. coli* core structures. When fit to proposed compositions,[3] the observed species were found to exhibit heterogeneity in heptose (Hep), hexose (Hex), 3-deoxy-D-*manno*-octulosonic acid (Kdo), phosphate, and phosphoethanolamine (PEA) (Table II). Specifically, two main core types were observed containing either 3 Hex and 3 Hep (with 2 or 3 Kdos) or 4 Hex and 4 Hep (with 2 Kdos), with variable amounts of phosphate and PEA in both. The pGEMLOS-7 *O*-deacylated LPS mixture contained many of these same species, in addition to two major new species at m/z 3334.5 and 3456.8 (Fig. 3). The m/z 3334.5 species apparently arises from the addition of Hex and *N*-acetylhexosamine (HexNAc) to the pGEM core structure containing 4 Hex, 4 Hep, 2 Kdo, 2 phosphates, and 1 *O*-deacylated diphosphorylated lipid A (*O*-DPLA) moiety. A further addition of 1 PEA moiety gives the m/z 3456.8 species. These data suggest that the transformation producing pGEMLOS-7 results in the addition of a Hex-HexNAc moiety to the *E. coli* LPS. Likewise, the main species in the pGEMLOS-5 *O*-deacylated LPS (m/z 3700.6 and 3823.6) were found to arise from the addition of 2 Hex plus 2 HexNAc to the pGEM core structure containing 4 Hex, 4 Hep, 2 Kdo, 2 phosphates, 1 *O*-DPLA, and 0 or 1 PEA. These structures are also found in the pGEMLOS-4 *O*-deacylated LPS, in addition to new species arising from the further addition of either another Hex (m/z 4083.2 and 4206.4) or HexNAc (m/z 4124.5 and 4246.8) to, in this case, the pGEM core structure containing 4 Hex, 4 Hep, 3 Kdo, 2 phosphates, 1 *O*-DPLA, and 0 or 1 PEA. Of the chimeric LPS structures, only these high molecular weight pGEMLOS-4 components contained the third Kdo moiety.

Analysis of Oligosaccharides by ESI-MS and LSIMS—The LPS from pGEM, pGEMLOS-4, pGEMLOS-5, and pGEM-

FIG. 4. **Negative ion ESI-MS of the oligosaccharide pools from the LPS of pGEM (*A*), pGEMLOS-7 (*B*), pGEMLOS-5 (*C*), and pGEMLOS-4 (*D*).** Peaks marked with *asterisks* are anhydro species.

[2] Ions measured by LSIMS and ESI-MS (unless otherwise noted) are monoisotopic masses of the ^{12}C-containing component, and ions observed by MALDI-TOF are average masses.

[3] Residue mass values (monoisotopic mass, average mass) for LOS and LPS structural moieties are as follows: Hex (162.053, 162.142), Hep (192.063, 192.169), HexNAc (203.079, 203.195), Kdo (220.058, 220.179), NeuAc (291.095, 291.258), phosphate (79.966, 79.980), PEA (123.008, 123.048), and diphosphorylated *O*-deacylated lipid A (934.457, 934.994).

LOS-7 were subjected to mild acid hydrolysis to liberate free oligosaccharides. Initially, small aliquots of the oligosaccharide fractions were desalted by size exclusion HPLC and analyzed as mixtures by negative ion ESI-MS (Fig. 4). The ESI-MS spectra contained predominantly doubly charged ions (M −

Chimeric LPS from E. coli

TABLE III

Molecular weights and proposed compositions of the oligosaccharides from pGEM, pGEMLOS-7, pGEMLOS-5, and pGEMLOS-4

Relative ion abundances (given in parentheses) represent the sum of the monoisotopic molecular ion and anhydro peaks for a given species (see Fig. 4). P indicates phosphate.

pGEM	pGEMLOS-7	pGEMLOS-5	pGEMLOS-4	Calculated M_r	Proposed compositions					
					HexNAc	Hex	Hep	P	PEA	Kdo
				Chimeric structures						
			2870.0 (12)	2870.8	3	6	4	2	1	1
			2828.6 (10)	2827.8	3	6	4	3		1
			2747.8 (13)	2747.8	3	6	4	2		1
			2706.1 (12)	2706.8	2	7	4	2		1
		2666.5 (20)	2666.7 (13)	2667.7	2	6	4	2	1	1
		2543.9 (58)	2544.0 (31)	2544.7	2	6	4	2		1
	2302.5 (30)			2302.6	1	5	4	2	1	1
	2177.7 (45)			2179.6	1	5	4	2		1
				Core structures						
	2058.6 (23)			2060.5		4	4	2	2	1
2016.7 (13)				2017.4		4	4	3	1	1
1937.2 (50)	1937.1 (52)	1937.1 (15)		1937.5		4	4	2	1	1
1893.5 (51)				1894.4		4	4	3		1
1814.0 (100)	1814.0 (54)	1814.0 (19)		1814.4		4	4	2		1
	1743.9 (19)			1745.4		4	3	2	1	1
	1705.8 (24)	1705.9 (11)	1705.8 (14)	1706.4		3	3	2	2	1
		1622.4 (15)		1622.4		4	3	2		1
1582.9 (49)	1582.8 (100)	1582.8 (53)	1582.9 (68)	1583.4		3	3	2	1	1
1539.2 (51)		1539.3 (16)	1539.3 (19)	1540.3		3	3	3		1
1459.3 (78)	1459.9 (50)	1459.9 (100)	1459.8 (100)	1460.3		3	3	2		1

$2H)^{2-}$. In general, the data were consistent with results from the MALDI-TOF analysis of *O*-deacylated LPS (Fig. 3). The pGEM sample was found to contain seven major oligosaccharides and several minor species, ranging in molecular weight from 1459.3 to 2016.7. As shown in Table III, proposed compositions were determined for the various species which indicated that the structures consisted of two main core types as follows: one containing 3 Hex, 3 Hep, and 1 Kdo, and another containing 4 Hex, 4 Hep, and 1 Kdo. Variability in the number of phosphate and PEA groups was responsible for the large number of species present in the mixture.

The pGEMLOS-4, pGEMLOS-5, and pGEMLOS-7 samples contained many of the species found in the pGEM sample, in addition to larger molecular weight oligosaccharides (Table III). New LPS glycoforms of M_r 2177.7 and 2302.5 were observed in the pGEMLOS-7 sample, consistent with the addition of a single Hex and HexNAc residue to the pGEM core structure containing 4 Hex, 4 Hep, 1 Kdo, 2 phosphates, and 0 or 1 PEA. The high molecular weight components of the pGEMLOS-5 sample (M_r 2543.9 and 2666.5) suggested the further addition of yet another Hex-HexNAc unit, and the pGEMLOS-4 sample contained even higher molecular weight materials (ranging from M_r 2706.1 to 2870.0) consistent with the addition of one more Hex or HexNAc moiety.

To aid in the determination of proposed compositions for these species, oligosaccharides from the pGEM, pGEMLOS-4, pGEMLOS-5, and pGEMLOS-7 samples were separated by size exclusion chromatography, and fractions were analyzed by LSIMS and/or ESI-MS. Selected fractions representing the two major pGEM core types and the various chimeric structures were then pooled, dephosphorylated with aqueous HF, rechromatographed on size exclusion HPLC, and analyzed again by negative ion LSIMS or ESI-MS. Proposed compositions for the molecular ions observed after HF treatment are listed in Table IV, and Fig. 5 shows examples of ESI spectra for four of the fractions. Upon removal of phosphate and PEA moieties, the major high mass species present in the pGEMLOS-7 sample is an oligosaccharide of average M_r 2020.3 (1 HexNAc, 5 Hex, 4 Hep, and 1 Kdo). The pGEMLOS-5 sample contains an oligosaccharide with an average M_r of 2386.3, resulting from the further addition of 1 Hex and 1 HexNAc to the pGEMLOS-7 LPS (2 HexNAc, 6 Hex, 4 Hep, and 1 Kdo). This species is also present in the pGEMLOS-4 sample, in addition to higher molecular weight structures containing an additional Hex (M_r average = 2548.4) or HexNAc (M_r average = 2589.5).

Monosaccharide Composition and Linkage Analyses—Mass spectrometric analyses of the free oligosaccharides from pGEM, pGEMLOS-4, pGEMLOS-5, and pGEMLOS-7 indicated that the different chimeric structures arise from additions of stoichiometric amounts of hexose and HexNAc residues to a variably phosphorylated pGEM core structure containing 4 Hex, 4 Hep, and 1 Kdo. No chimeric structures were observed to contain the 3 Hex, 3 Hep, and 1 Kdo core. For comparison purposes, dephosphorylated oligosaccharide fractions containing the two pGEM core types and the main chimeric structures from pGEMLOS-4, pGEMLOS-5, and pGEMLOS-7 were hydrolyzed in 2 N trifluoroacetic acid to determine their monosaccharide compositions and therefore the identities of the Hex and HexNAc residues. When analyzed by high pH anion exchange chromatography with pulsed amperometric detection, the pGEM hydrolysates were found to contain only galactose, glucose, and L-*glycero*-D-*manno*-heptose (Table V). (The Kdo residue is not recovered under these hydrolysis conditions.) The two core types were identified as $GalGlc_2Hep_3$ and $GalGlc_3Hep_4$. The pGEMLOS-7 sample contained $GlcNH_2$ $Gal_2Glc_3Hep_4$, suggesting that the larger pGEM core was being modified by the addition of one Gal and one *N*-acetylglucosamine (GlcNAc) residue. Likewise, the composition of the pGEMLOS-5 sample suggested that the larger pGEM core was being further glycosylated with only Gal and GlcNAc residues. Fraction 2 from pGEMLOS-4, which contained the same species as pGEMLOS-5, gave similar results, and fraction 1 from pGEMLOS-4, which contains three main species (see Table IV), contained slightly more $GlcNH_2$.

Aliquots of the same six dephosphorylated oligosaccharide fractions used for monosaccharide composition analysis were taken for methylation analysis to establish sugar linkage positions. The partially methylated alditol acetates observed by gas chromatography/MS are listed in Table VI. Again, by compar-

TABLE IV

Molecular weights and proposed compositions of the dephosphorylated oligosaccharides from pGEM, pGEMLOS-7, pGEMLOS-5, and pGEMLOS-4

| Fraction	Observed M_r (average)	Calculated M_r		Proposed compositions			
		Average	Exact	HexNAc	Hex	Hep	Kdo
pGEM (fraction 2)	1300.8	1301.1	1300.4		3	3	1
pGEM (fraction 1)	1655.3	1655.4	1654.5		4	4	1
pGEMLOS-7	2020.3	2020.2	2019.7	1	5	4	1
pGEMLOS-5	2386.3	2386.1	2384.8	2	6	4	1
pGEMLOS-4 (fraction 2)	2386.3	2386.1	2384.8	2	6	4	1
pGEMLOS-4 (fraction 1)	2386.1	2386.1	2384.8	2	6	4	1
	2548.4	2548.2	2546.8	2	7	4	1
	2589.5	2589.3	2587.9	3	6	4	1
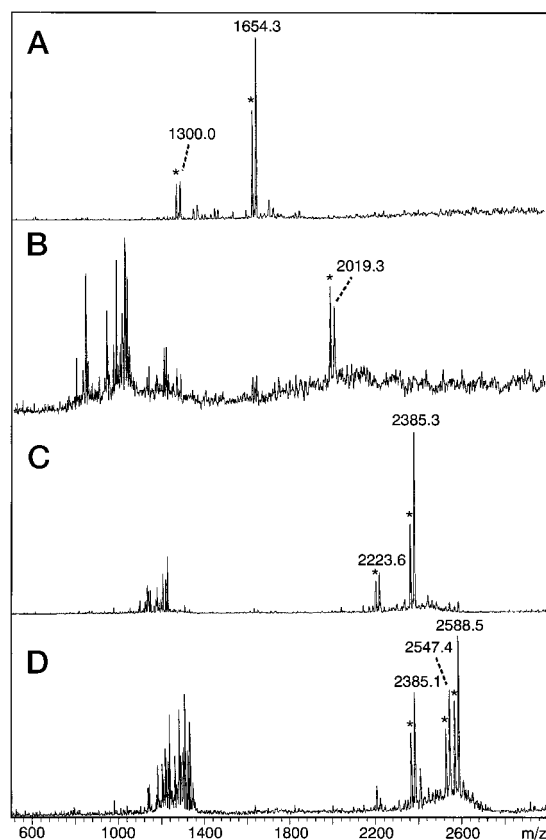
FIG. 5. **Negative ion ESI-MS of the dephosphorylated oligosaccharide fractions used for composition and methylation analyses: pGEM fraction 1 (*A*), pGEMLOS-7 (*B*), pGEMLOS-5 (*C*), and pGEMLOS-4 fraction 1 (*D*).** Ions below m/z 1400 in spectra *B–D* are doubly charged ions arising from incompletely dephosphorylated glycoforms.

ing the two pGEM core types, it is relatively straightforward to see that the second terminal heptose of the larger pGEM core is converted to a 1,7-linked heptose in all of the chimeric structures and thus must represent the linkage site for the novel glycosylation. Since no chimeric structures were observed with the Hep₃ core, it is most likely that the non-reducing terminal heptose recently identified on the oligosaccharide branch in the K-12 core structure (18) is the modified terminal heptose. Additionally, no new tri-linked saccharides were obtained from the chimeric oligosaccharides, suggesting that the sugars were most likely all added in a straight chain.

Sequencing of Chimeric Oligosaccharides by MS/MS—To confirm the identity of the linkage site between the *E. coli* LPS core and the novel oligosaccharide moieties, and to determine the sequences of the added sugars, the dephosphorylated oligosaccharides were subjected to MS/MS analysis. For these

experiments, samples were run in the positive ion mode and doubly charged molecular ions, $(M + 2H)^{2+}$, were selected for collision-induced dissociation (CID). The resulting MS/MS spectra of the pGEM core oligosaccharide (M_r 1654.5) and the pGEMLOS-5 chimeric structure (M_r 2384.8) are shown in Fig. 6. Various reducing terminal (Y-type) and non-reducing terminal (B-type) sequence ions are present in the spectra. For the pGEM oligosaccharide (Fig. 6*A*), the Y ion series including the $Y_{6\alpha'}$ (m/z 732.2 (2+)), $Y_{5\alpha'}$ (m/z 651.2 (2+)), and $Y_{4\alpha'}$ (m/z 1139.3) fragment ions and the corresponding B ion series including the $B_{3\alpha'}$ (m/z 517.2), $B_{4\alpha'}$ (m/z 841.3), B_5 (m/z 1225.4), and B_6 (m/z 1417.4) fragment ions support the published structure with the fourth heptose on the non-reducing terminus of the largest oligosaccharide branch. In addition to these sequence ions, several ions present in the spectrum apparently arise from internal cleavages, which can occur under high energy CID conditions. In the spectrum of the pGEMLOS-5 oligosaccharide (Fig. 6*B*), two similar Y and B type ion series clearly define the sequence and linkage site of the added tetrasaccharide. Intense B ions at m/z 366.1 ($B_{2\alpha'}$) and 731.3 ($B_{4\alpha'}$) arise from the sequential cleavage of two Hex-HexNAc moieties. These losses are also represented by the corresponding $Y_{9\alpha'}$ (m/z 2020.6) and $Y_{7\alpha'}$ (m/z 1655.5) fragment ions. Fragment ions at m/z 923.3 ($B_{5\alpha'}$) and m/z 1463.4 ($Y_{6\alpha'}$) confirm that the Hex-HexNAc-Hex-HexNAc moiety is linked to a heptose, and additional cleavages further along the large oligosaccharide branch (the $B_{6\alpha'}$, $B_{7\alpha'}$, and $Y_{5\alpha'}$ ions) confirm that the novel tetrasaccharide is attached to the largest branch of the pGEM core structure, as indicated in Fig. 6*B*.

In the MS/MS spectra of the chimeric oligosaccharides from pGEMLOS-7 and pGEMLOS-4, intense B ions also clearly defined the structures of the added sugar moieties. In the pGEMLOS-7 oligosaccharide (M_r 2019.7), a B ion at m/z 366.1 corresponds to a single Hex-HexNAc disaccharide moiety. The pGEMLOS-4 oligosaccharide of M_r 2587.9 (HexNAc₃Hex₆Hep₄Kdo) lost a HexNAc-Hex-HexNAc fragment (m/z 569.2) and a HexNAc-Hex-HexNAc-Hex-HexNAc fragment (m/z 934.3), whereas the pGEMLOS-4 oligosaccharide of M_r 2546.8 (HexNAc₂Hex₇Hep₄Kdo) lost a Hex-Hex-HexNAc (m/z 528.2) and a Hex-Hex-HexNAc-Hex-HexNAc (m/z 893.3) fragment. In addition to those B-type ions, the latter spectrum also contained large ions at m/z 366.1 and 731.3, which apparently arise as internal fragments in that case.

Assuming that the oligosaccharides are built up sequentially, *i.e.* from pGEMLOS-7 to pGEMLOS-5 to pGEMLOS-4, the MS/MS data, in combination with our methylation analysis results, allows the partial structures of the chimeric oligosaccharides to be deduced as shown in Fig. 7.

DISCUSSION

In this study, we have provided structural data to support our original hypothesis that *E. coli* K-12 transformed with plasmids containing portions of an eight gene segment from *H.*

TABLE V
Monosaccharide compositions of the dephosphorylated oligosaccharide fractions

Molar ratios were derived from comparison to the hydrolysate of the *S. typhimurium* Ra oligosaccharide of known composition, and then those values were normalized to either 3.0 or 4.0 heptoses in each fraction.

	pGEM (fraction 2)	pGEM (fraction 1)	pGEMLOS-7	pGEMLOS-5	pGEMLOS-4 (fraction 2)	pGEMLOS-4 (fraction 1)
GlcN			1.0	2.7	2.4	3.3
Gal	0.9	1.1	1.7	3.0	2.8	3.1
Glc	2.2	3.2	3.2	3.4	3.4	3.4
Hep	3.0	4.0	4.0	4.0	4.0	4.0

TABLE VI
Methylation analysis of the dephosphorylated oligosaccharide fractions

Peak areas were measured from the GC/MS EI total ion chromatograms, and values were normalized to the 1,3,6-Glc residue. The data for pGEMLOS-7 are the average of two runs.

	pGEM (fraction 2)	pGEM (fraction 1)	pGEMLOS-7	pGEMLOS-5	pGEMLOS-4 (fraction 2)	pGEMLOS-4 (fraction 1)
T-Glc	0.5	0.2				
T-Gal	0.6	0.5	1.2	0.9	0.7	0.7
1,2-Glc	0.2	0.7	1.0	0.8	0.8	0.7
1,3-Gal				0.4	0.5	0.6
1,6-Glc	0.4	1.1	1.6	1.5	1.3	1.3
1,6-Gal					0.3	0.6
1,3,6-Glc	1.0	1.0	1.0	1.0	1.0	1.0
T-Hep	0.3	0.6	0.6	0.3	0.3	0.2
1,3-Hep	0.3	0.3	0.3	0.4	0.6	0.4
1,7-Hep			0.4	0.6	0.5	0.6
1,3,7-Hep	0.4	0.4	0.2	0.3	0.5	0.5
T-GlcNAc			0.1	0.2	0.2	0.3
1,4-GlcNAc				0.9	1.1	1.0
1,3-GlcNAc			0.3	0.7	0.7	1.0

influenzae involved in LOS biosynthesis makes chimeric LPS (10). Moreover, we have shown that the chimeric LPS are segregated hybrid-type structures, where the *E. coli* r-LPS core structure is first synthesized and then serves as a scaffold for *H. influenzae* LOS biosynthesis enzymes to add a second independent set of sugars not found in the parent *E. coli* strain. Thus, the biosynthetic pathways appear to be sequential (segregated) and not intermixed.

One of the more interesting outcomes of the chemical analysis of these chimeric LPS species is the role of the terminal branch heptose in the *E. coli* r-LPS as the acceptor for oligosaccharide elongation. When we began this work, the published structure for the complete *E. coli* K-12 core did not contain a second terminal heptose but rather had this fourth heptose as part of the inner core region (31). The oligosaccharide branch was believed to terminate in glucose, which was proposed to be the acceptor site for O-antigen and other substituents (31). From our initial observations, it was apparent that only *E. coli* r-LPS structures containing this fourth heptose (*i.e.* complete core structures) underwent elongation in the plasmid-transformed chimeric strains. When a correction to the complete K-12 structure was published that placed the fourth heptose on the non-reducing terminus of the oligosaccharide branch (18), as opposed to the inner core, our findings became easier to interpret. The revised core structure is consistent with the linkage analysis data obtained in our study and also provides a more accessible and logical acceptor site for chain elongation.

In the chimeric structures, GlcNAc is the first sugar added to the 7-position of this heptose. We found that for this crucial first step in the elongation sequence to occur, *E. coli* must have a functional *wecA* gene. WecA is involved in O-antigen biosynthesis in *E. coli*. There are several different pathways involved in O-antigen biosynthesis. In the biosynthesis pathways utilized by homopolymeric O-antigens in *E. coli* 08 and 09, the polymer is elongated on the non-reducing terminus of

a GlcNAc-pyrophosporyl-undecaprenol carrier lipid (19, 32, 33). This pathway is initiated by WecA that transfers UDP-GlcNAc to the isoprenoid carrier. The GlcNAc subsequently accepts sequentially transferred monosaccharides one unit at a time. The GlcNAc is used to initiate the polymer and becomes part of the repeating O-antigen unit in *E. coli* 018, 075, and 0111 (32). In the case of *E. coli* K-12, it was recently shown that when mutations causing the rough phenotype are complemented, the complemented strains can also produce an O-antigen which has GlcNAc at the reducing terminus of the repeat unit (34). The inability to form an *H. influenzae–E. coli* LOS chimeric structure in an *E. coli* wecA− strain 21548 indicates that *wecA* is required for synthesis of *H. influenzae-E. coli* LPS chimeric structure. It appears that the GlcNAc not only accepts the chimeric *H. influenzae* structure but is incorporated into the final chimeric structure. Our studies have shown that *lsgG* is also necessary for the formation of the chimeric structures since if it is absent from the plasmid constructs with all other genes of the *lsg* cluster, the chimeric LPS is not formed (16). LsgG has been shown to be a regulatory gene in the *modABCD* operon and has high homology to *modE*, the gene which regulates this operon (35) (Table VII). It is interesting to speculate that *lsgG* may be also involved in the regulation of *wecA*, enhancing the efficiency of the chimeric LPS formation. We are currently performing studies to test this hypothesis.

The second step in the biosynthesis of the chimeric LPS is the addition of galactose to the 3-position of the terminal GlcNAc. The resulting disaccharide, Gal1→3GlcNAc1→, is the structural moiety observed in pGEMLOS-7, which arises when the transforming plasmid contains *lsgF* and *lsgG* from *Haemophilus*. Examination of the predicted amino acid sequence of LsgF indicates that it has high homology to a galactosyltransferase (AmsD) from *Erwinia amylovora* (36), suggesting that *lsgF* may encode a galactosyltransferase in *Haemophilus*, which evidently is capable of recognizing GlcNAc on the pyro-
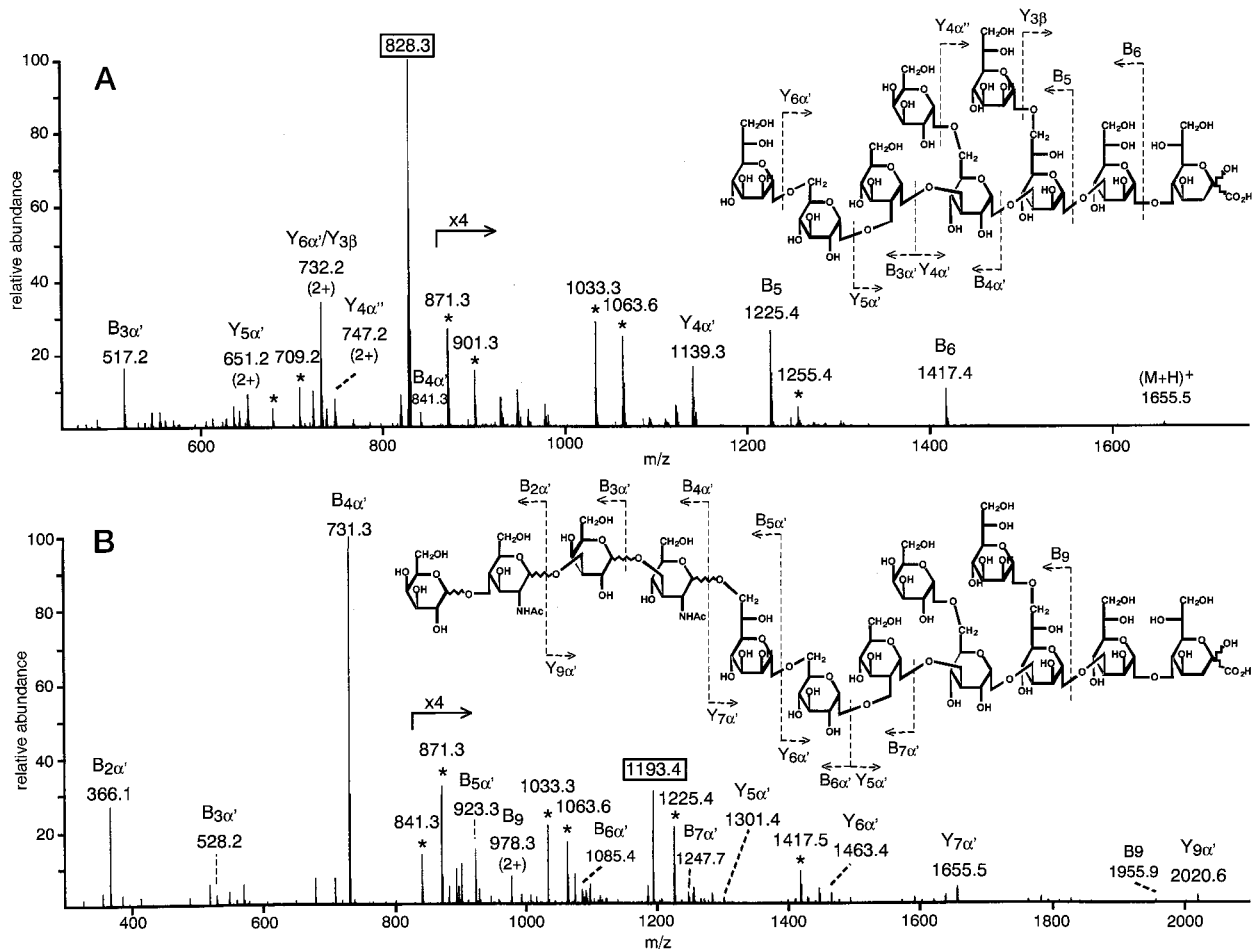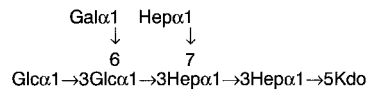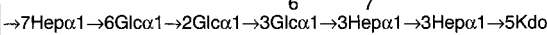
FIG. 6. **MS/MS spectra of the dephosphorylated oligosaccharides from pGEM (*A*) and pGEMLOS-5 (*B*).** The doubly charged parent ions are indicated with *boxes*. Peaks marked with *asterisks* are believed to be internal fragments, arising from multiple bond cleavages.
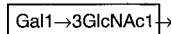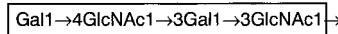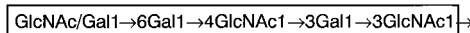

FIG. 7. **Proposed structures of the chimeric oligosaccharides.** Only the complete *E. coli* K-12 core structure containing a fourth heptose on the terminus of the oligosaccharide branch undergoes modification. Additional saccharides (designated *R*) are added to the 7-position of this heptose to form the chimeric oligosaccharides.

phosporyl-undecaprenol carrier lipid as a substrate. Although unmodified *E. coli* core LPS structures remain in all of the chimeric LPS mixtures, the lack of intermediate chimeric structures as one progresses from pGEMLOS-7 to pGEMLOS-5 and pGEMLOS-4 suggest that once galactose is added to the GlcNAc-undecaprenyl phosphate, the addition of the other *Haemophilus*-related sugars proceeds to defined end points.

Therefore, the nonstoichiometric glycosylation of the *E. coli* core LPS by the *Haemophilus* oligosaccharides may be a function of the efficiency of the ligase that transfers O-antigen to the LPS core. This would be analogous to Enterobacteria O-antigen biosynthesis where one also observes a significant population of unmodified core LPS.

In the pGEMLOS-5 strain, an additional three genes are contained in the transforming plasmid (*lsgC-lsgG*), and an additional GlcNAc and Gal are observed in the resulting LPS. These sugars now define the tetrasaccharide Gal1→4GlcNAc1→3Gal1→3GlcNAc1→. The LPS from this transformant is now reactive to the 3F11 mAb, suggesting that this new disaccharide is β-linked to form the terminal trisaccharide, Galβ1→4GlcNAcβ1→3Gal1→. All of the new *lsg* genes contained in this plasmid have some homology with known glycosyltransferase genes as follows: *lsgC* has homology with the *amsD* from *E. amylovora*, which encodes a galactosyltransferase (36), and *trsD* from *Yersinia enterocolitica*, a gene involved in LPS inner core synthesis (37); *lsgD* has homology with *lgtE* from *Neisseria gonorrhoeae* (38) and *lgtB* from *N. gonorrhoeae* and *Neisseria meningitidis* (39); and *lsgE* has homology with a putative glycosyltransferase gene from *Actinobacillus* sp. (40), the galactosyltransferase gene *amsB* from *E. amylovora* (36), and the glycosyltransferase gene *lgtD* from *Rickettsia prowazekii* (41). The fact that these three additional *lsg* genes in the transforming plasmid apparently result in the addition of only two more sugars to the growing oligosaccharide chain may indicate that the acceptor for one of the glycosyltransferases is absent in the chimeric LPS.

When two more *lsg* genes are added in the transforming

Chimeric LPS from E. coli

TABLE VII
Gene homology

Lsg	Homologous genes/proteins	Ref.	Protein homology	Protein function	Proposed function(s) of *Haemophilus* gene product
LsgA	Wzx from *E. coli*	43	22% identity 44% positive	Putative O-antigen transporter	Flippase?[a]
	TrsA from *Y. enterocolitica*	44	24% identity 44% positive	Putative O-antigen transporter	
	Wzx from *Bacteriodes fragilis*	49	20% identity 36% positive	Putative O-antigen flippase	
LsgB	Sialyltransferase gene from *N. meningitidis*	42	27% identity 46% positive	α-2,3-Sialyltransferase	Sialyltransferase or other glycosyltransferase[b]
LsgC	AmsD from *E. amylovora*	36	26% identity 47% positive	Galactosyltransferase	1,4-Galactosyltransferase
	TrsD from *Y. enterocolitica*	37	38% identity 52% positive	Involved in LPS inner core synthesis	
LsgD	LgtE from *N. gonorrhoeae*	38	22% identity 40% positive	Lacto-*N*-neotetraose glycosyltransferase	1,3-*N*-Acetylglucosaminyltransferase[c]
	LgtB from *N. gonorrhoeae* and *N. meningitidis*	39	24% identity 41% positive	Lacto-*N*-neotetraose glycosyltransferase	
LsgE	Putative glycosyltransferase gene from *Actinobacillus* sp.	40	77% identity 84% positive	Putative glycosyltransferase	Galactosyltransferase[d]
	AmsB from *E. amylovora*	36	27% identity 49% positive	Galactosyltransferase	
	LgtD from *R. prowazekii*	41	23% identity 47% positive	Glycosyltransferase	
LsgF	AmsD from *E. amylovora*	36	45% identity 66% positive	Galactosyltransferase	1,3-Galactosyltransferase
	Putative deoxy-L-talan synthesis gene from *Actinobacillus* sp.	50	83% identity 90% positive	Glycosyltransferase	
	GDP-mannose biosynthesis gene from *E. coli*	51	37% identity 54% positive	Glycosyltransferase	
LsgG	ModE from *E. coli*	35	47% identity 64% positive	Molybdenum transport and regulation of transcription	Regulation or modulation of the *E. coli* *wecA* gene, which results in the addition of GlcNAc to the branch heptose
	MopB from *Rhodobacter capsulatus*	52	32% identity 51% positive	Molybdenum-pterin-binding protein	

[a] When *lsgA* and *lsgB* are both present, the pGEMLOS-5 tetrasaccharide structure can be extended with either a 6-linked galactose or GlcNAc to form the structures seen in pGEMLOS-4 LPS, which also contain a third Kdo moiety which may be the epitope recognized by mAb 6E4. Thus, the functions of these two *Haemophilus* ORFs remain unclear.

[b] Recent studies with a double gene knockout of *lst*, a gene with high homology to the sialyltransferase in *Haemophilus ducreyi* (53), and *lsgB* suggest that *lsgB* may also encode a sialyltransferase.

[c] The structure of the pGEMLOS-5 LPS suggests that one of the three ORFs C, D, or E encodes a 1,3-*N*-acetylglucosaminyltransferase and another encodes a 1,4-galactosyltransferase. Based on our previous investigations of the LOS formed when *lsgD* or *lsgE* are knocked out in mutant strains 281.25 and 276.4, respectively (6), we can conclude that *lsgD* is likely to encode a 1,3-*N*-acetylglucosaminlyltransferase and *lsgC* and *lsgE* are both likely to encode galactosyltransferases.

[d] The LOS formed when *lsgE* is knocked out (6) suggest that *lsgE* encodes a galactosyltransferase whose acceptor may be absent in the chimeric LPS.

plasmid (*lsgA-lsgG*) to form the pGEMLOS-4 chimeric strain, we observe that the 3F11 epitope disappears, and the terminal Gal residue of the epitope is capped by either a second Gal or a GlcNAc moiety, apparently linked to the 6-position of the Gal. These new species present in the pGEMLOS-4 LPS population were also observed to contain a third Kdo moiety, presumably somewhere in their core regions. Whereas some of the incomplete core structures found in the wild-type *E. coli* K-12 LPS populations also contain a third Kdo, of the chimeric structures, only the structures unique to pGEMLOS-4 were found to contain a third Kdo. This chimeric strain was also recognized by mAb 6E4 which recognizes an inner core, Kdo-related epitope in *H. influenzae*, suggesting that this third Kdo forms a different epitope than the one found in the core structure of the wild-type *E. coli* LPS. Thus, the addition of *lsgA* and *lsgB* to the transforming plasmid which formed strain pGEMLOS-4 seems to have multiple effects on the chimeric LPS structure. *lsgB* is homologous to the sialyltransferase gene from *N. meningitidis* (42). *lsgA* is homologous to both the *wzx* gene product from *E. coli* (43) and TrsA of *Y. enterocolitica* (44). These are putative O-antigen transporters, suggesting that *lsgA* may encode a flippase.

Whereas sialyl-*N*-acetyllactosamine-containing structures are only minor components of the wild-type *H. influenzae* type b strain A2 LOS population (5), we have previously seen that *lsg* genes are involved in the synthesis of this epitope. Transposon mutagenesis of *lsgD* produced mutant strain 281.25, which lost all ability to add galactose to Hib LOS glycoforms (6). This strain could not make any of the wild-type LOS structures larger than the major species containing four glucoses and three heptoses. Mutation of *lsgE* (which is downstream of *lsgD*) produced strain 276.4 which had essentially the same defect, except for one important difference, strain 276.4 retained the ability to make the sialyl-*N*-acetyllactosamine epitope. These results suggest that in the transposon mutants, the knockout of *lsgD* has a polar effect on *lsgE*, which would imply that the gene product of *lsgE* is a galactosyltransferase required for synthesis of the higher molecular weight wild-type structures containing terminal galactose(s) on their glucose disaccharide branches, and the gene product of *lsgD* is likely an *N*-acetylglucosaminyltransferase required for the synthesis of the sialyl-*N*-acetyllactosamine epitope (see Table VII). The case for these assignments can be made on the basis of the homologies noted above (*lsgE* is homologous to a galactosyltransferase gene from *E. amylovora*) and the LOS glycoforms observed in the 276.4 and 281.25 mutant strains. Since no

truncated versions of the sialyl-*N*-acetyllactosamine structure were seen in the 276.4 LOS population (*i.e.* no species lacking either sialic acid or sialic acid plus galactose), *lsgD* may encode the glycosyltransferase which adds the GlcNAc to the oligosaccharide branch. This is also consistent with the observation that one of the genes in *lsgC-lsgE* is apparently responsible for adding GlcNAc to the 3-position of the Gal which is terminal in the pGEMLOS-7 LPS structure.

CONCLUSIONS

In this study, we have transfected a strain of *E. coli* that normally synthesizes a rough LPS with a plasmid containing a variable set of exogenous genes known to be involved in the synthesis of carbohydrate epitopes specific to *H. influenzae* (16). The resulting chimeric LPS have been shown to contain additional sugars that have been previously identified as terminal components of native *H. influenzae* lipooligosaccharides (5). This chimeric carbohydrate expression system has provided information that is relevant to unraveling the functions of these *lsg* genes and has the additional advantage of being carried out in the absence of the normal endogenous genetic background on *H. influenzae*. Indeed, while gene knockouts of some of the *lsg* genes in *H. influenzae* have been completed (6), downstream or regulatory gene effects can often complicate their functional analysis. In this *E. coli* expression system, structural analysis of the resulting chimeric LPS has shown that synthesis proceeded as a serial (non-parallel) synthesis, that is the new elements of the chimeric LPS were added after the formation of the *E. coli* r-LPS. The fact that this synthesis was sequential (rather than interdigitated with the r-LPS synthesis, for example) allowed for the functions of these *H. influenzae* gene products to be more readily delineated from the chimeric oligosaccharide structures. Moreover, screening of the chimeric LPS products with monoclonal antibodies such as 3F11 enabled us to follow the formation of terminal sugar sequences (epitopes) that are unique to the *Haemophilus* strain from which the plasmid DNA originated.

In addition to its use in unraveling carbohydrate-specific gene functions, this methodology has potential as a general method for the production of specific oligosaccharides or as a method for the production of carbohydrate-based vaccine components. For example, the high (and commercially viable) output of LPS biosynthesis in *E. coli* coupled with its transformation with exogenous oligosaccharide-specific biosynthesis genes could prove to be a relatively inexpensive and high throughput approach for the synthesis of large amounts of complex carbohydrates. Indeed, our largest chimera contained a novel pentasaccharide attached to the host *E. coli* r-LPS acceptor, although one could easily imagine further extensions to even larger and more complex structures. One particularly exciting application of this methodology might be in the construction of carbohydrate-based vaccines, where carbohydrate epitopes from pathogenic bacteria could be incorporated into non-lethal strains. Likewise, the resulting chimeric LPS could be made to be non-toxic as well, such as through modification of the lipid A moiety through gene manipulation (45, 46), yielding novel non-toxic LPS bearing specific epitopes.

REFERENCES

1. Kimura, A., and Hansen, E. J. (1986) *Infect. Immun.* **51,** 69–79
2. Kimura, A., Patrick, C. C., Miller, E. E., Cope, L. D., McCracken, G. H., Jr., and Hansen, E. J. (1987) *Infect. Immun.* **55,** 1979–1986
3. Zwahlen, A., Rubin, L. G., and Moxon, E. R. (1986) *Microb. Pathog.* **1,** 465–473
4. Masoud, H., Moxon, E. R., Martin, A., Krajcarski, D., and Richards, J. C. (1997) *Biochemistry* **36,** 2091–2103
5. Phillips, N. J., Apicella, M. A., Griffiss, J. M., and Gibson, B. W. (1993) *Biochemistry* **32,** 2003–2012
6. Phillips, N. J., McLaughlin, R., Miller, T. J., Apicella, M. A., and Gibson, B. W. (1996) *Biochemistry* **35,** 5937–5947
7. Schweda, E. K. H., Hegedus, O. E., Borrelli, S., Lindberg, A. A., Weiser, J. N., Maskell, D. J., and Moxon, E. R. (1993) *Carbohydr. Res.* **246,** 319–330
8. Schweda, E. K. H., Jansson, P.-E., Moxon, E. R., and Lindberg, A. A. (1995) *Carbohydr. Res.* **272,** 213–224
9. Weiser, J. N. (1992) *Microb. Pathog.* **13,** 335–342
10. Spinola, S. M., Abu Kwaik, Y., Lesse, A. J., Campagnari, A. A., and Apicella, M. A. (1990) *Infect. Immun.* **58,** 1558–1564
11. McLaughlin, R., Lee, N.-G., Abu Kwaik, Y., Spinola, S. M., and Apicella, M. A. (1994) *J. Endotoxin Res.* **1,** 165–174
12. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) *Science* **269,** 496–512
13. Hood, D. W., Deadman, M. E., Allen, T., Masoud, H., Martin, A., Brisson, J. R., Fleischmann, R., Venter, J. C., Richards, J. C., and Moxon, E. R. (1996) *Mol. Microbiol.* **22,** 951–965
14. McLaughlin, R., Spinola, S. M., and Apicella, M. A. (1992) *J. Bacteriol.* **174,** 6455–6459
15. Mandrell, R. E., McLaughlin, R., Abu Kwaik, Y., Lesse, A., Yamasaki, R., Gibson, B., Spinola, S. M., and Apicella, M. A. (1992) *Infect. Immun.* **60,** 1322–1328
16. Abu Kwaik, Y., McLaughlin, R. E., Apicella, M. A., and Spinola, S. M. (1991) *Mol. Microbiol.* **5,** 2475–2480
17. Phillips, N. J., Apicella, M. A., Griffiss, J. M., and Gibson, B. W. (1992) *Biochemistry* **31,** 4515–4526
18. Holst, O., Zähringer, U., Brade, H., and Zamojski, A. (1991) *Carbohydr. Res.* **215,** 323–335
19. Meier-Dieter, U., Barr, K., Starman, R., Hatch, L., and Rick, P. D. (1992) *J. Biol. Chem.* **267,** 746–753
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21. Darveau, R. P., and Hancock, R. E. W. (1983) *J. Bacteriol.* **155,** 831–838
22. Tsai, C. M., and Frasch, C. E. (1982) *Anal. Biochem.* **119,** 115–119
23. Helander, I. M., Lindner, B., Brade, H., Altmann, K., Lindberg, A. A., Rietschel, E. T., and Zähringer, U. (1988) *Eur. J. Biochem.* **177,** 483–492
24. Hardy, M. R., Townsend, R. R., and Lee, Y. C. (1988) *Anal. Biochem.* **170,** 54–62
25. Levery, S. B., and Hakomori, S.-I. (1987) *Methods Enzymol.* **138,** 13–25
26. Larson, G., Karlsson, H., Hansson, G. C., and Pimlott, W. (1987) *Carbohydr. Res.* **161,** 281–290
27. Phillips, N. J., John, C. M., Reinders, L. G., Gibson, B. W., Apicella, M. A., and Griffiss, J. M. (1990) *Biomed. Environ. Mass Spectrom.* **19,** 731–745
28. Falick, A. M., Wang, G. H., and Walls, F. C. (1986) *Anal. Chem.* **58,** 1308–1311
29. Bateman, R. H., Green, M. R., Scott, G., and Clayton, E. (1995) *Rapid Commun. Mass Spectrom.* **9,** 1227
30. Gibson, B. W., Engstrom, J. J., John, C. M., Hines, W., and Falick, A. M. (1997) *J. Am. Soc. Mass Spectrom.* **8,** 645–658
31. Prehm, P., Schmidt, G., Jann, B., and Jann, K. (1976) *Eur. J. Biochem.* **70,** 171–177
32. Alexander, D. C., and Valvano, M. A. (1994) *J. Bacteriol.* **176,** 7079–7084
33. Klena, J. D., and Schnaitman, C. A. (1993) *Mol. Microbiol.* **9,** 393–402
34. Stevenson, G., Neal, B., Liu, D., Hobbs, M., Packer, N. H., Batley, M., Redmond, J. W., Lindquist, L., and Reeves, P. (1994) *J. Bacteriol.* **176,** 4144–4156
35. Grunden, A. M., Ray, R. M., Rosentel, J. K., Healy, R. G., and Shanmugam, K. T. (1996) *J. Bacteriol.* **178,** 735–744
36. Bugert, P., and Geider, K. (1995) *Mol. Microbiol.* **15,** 917–933
37. Skurnik, M., Venho, R., Toivanen, P., and al-Hendy, A. (1995) *Mol. Microbiol.* **17,** 575–594
38. Gotschlich, E. C. (1994) *J. Exp. Med.* **180,** 2181–2190
39. Jennings, M. P., Hood, D. W., Peak, I. R., Virji, M., and Moxon, E. R. (1995) *Mol. Microbiol.* **18,** 729–740
40. Yoshida, Y., Nakano, Y., Yamashita, Y., and Koga, T. (1998) *Infect. Immun.* **66,** 107–114
41. Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Ponten, T., Alsmark, U. C., Podowski, R. M., Naslund, A. K., Eriksson, A. S., Winkler, H. H., and Kurland, C. G. (1998) *Nature* **396,** 133–140
42. Gilbert, M., Watson, D. C., Cunningham, A. M., Jennings, M. P., Young, N. M., and Wakarchuk, W. W. (1996) *J. Biol. Chem.* **271,** 28271–28276
43. Yao, Z., and Valvano, M. A. (1994) *J. Bacteriol.* **176,** 4144–4156
44. Zhang, L., Radziejewska-Lebrecht, J., Krajewska-Pietrasik, D., Toivanen, P., and Skurnik, M. (1997) *Mol. Microbiol.* **23,** 63–76
45. Sunshine, M. G., Gibson, B. W., Engstrom, J. J., Nichols, W. A., Jones, B. D., and Apicella, M. A. (1997) *J. Bacteriol.* **179,** 5521–5533
46. Lee, N.-G., Sunshine, M. G., Engstrom, J. J., Gibson, B. W., and Apicella, M. A. (1995) *J. Biol. Chem.* **270,** 27151–27159

Chimeric LPS from E. coli

47. Schneider, H., Griffiss, J. M., Boslego, J. W., Hitchcock, P. J., Zahos, K. M., and Apicella, M. A. (1991) *J. Exp. Med.* **174,** 1601–1605
48. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene* (*Amst.*) **33,** 103–109
49. Comstock, L. E., Coyne, M. J., Tzianabos, A. O., Pantosti, A., Onderdonk, A. B., and Kasper, D. L. (1999) *Infect. Immun.* **67,** 3525–3532
50. Nakano, Y., Yoshida, Y., Yamashita, Y., and Koga, T. (1998) *Biochim. Biophys. Acta* **1442,** 409–414
51. Marolda, C. L., and Valvano, M. A. (1995) *J. Bacteriol.* **177,** 5539–5546
52. Wang, G., Angermuller, S., and Klipp, W. (1993) *J. Bacteriol.* **175,** 3031–3042
53. Bozue, J. A., Tullius, M. V., Wang, J., Gibson, B. W., and Munson, R. S., Jr. (1999) *J. Biol. Chem.* **274,** 4106–4114

Characterization of Chimeric Lipopolysaccharides from *Escherichia coli* Strain JM109 Transformed with Lipooligosaccharide Synthesis Genes (*lsg*) from *Haemophilus influenzae*

Nancy J. Phillips, Theresa J. Miller, Jeffrey J. Engstrom, William Melaugh, Robert McLaughlin, Michael A. Apicella and Bradford W. Gibson

Access the most updated version of this article at http://www.jbc.org/content/275/7/4747

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 52 references, 22 of which can be accessed free at
http://www.jbc.org/content/275/7/4747.full.html#ref-list-1