INFECTION AND IMMUNITY, Mar. 1995, p. 818–824
0019-9567/95/$04.00+0

Vol. 63, No. 3

Molecular Cloning and Characterization of the Nontypeable *Haemophilus influenzae* 2019 *rfaE* Gene Required for Lipopolysaccharide Biosynthesis

NA-GYONG LEE, MELVIN G. SUNSHINE, AND MICHAEL A. APICELLA*

Department of Microbiology, University of Iowa, Iowa City, Iowa

Received 23 September 1994/Returned for modification 22 November 1994/Accepted 12 December 1994

The lipooligosaccharide (LOS) of nontypeable *Haemophilus influenzae* (NTHi) is an important factor in pathogenesis and virulence. In an attempt to elucidate the genes involved in LOS biosynthesis, we have cloned the *rfaE* gene from NTHi 2019 by complementing a *Salmonella typhimurium rfaE* mutant strain with an NTHi 2019 plasmid library. The *rfaE* mutant synthesizes lipopolysaccharide (LPS) lacking heptose, and the *rfaE* gene is postulated to be involved in ADP-heptose synthesis. Retransformation with the plasmid containing 4 kb of NTHi DNA isolated from a reconstituted mutant into *rfaE* mutants gave wild-type LPS phenotypes. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis confirmed the conversion of the *rfaE* mutant LPS to a wild-type LPS phenotype. Sequence analysis of a 2.4-kb *Bgl*II fragment revealed two open reading frames. One open reading frame encodes the RfaE protein with a molecular weight of 37.6 kDa, which was confirmed by in vitro transcription and translation, and the other encodes a polypeptide highly homologous to the *Escherichia coli* HtrB protein. These two genes are transcribed from the same promoter region into opposite directions. Primer extension analysis of the *rfaE* gene revealed a single transcription start site at 37 bp upstream of the predicted translation start site. The upstream promoter region contained a sequence (TA AAAT) homologous to the −10 region of the bacterial σ^{70}-dependent promoters at an appropriate distance (7 bp), but no sequence resembling the consensus sequence of the −35 region was found. These studies demonstrate the ability to use complementation of defined LPS defects in members of the family *Enterobacteriaceae* to identify LOS synthesis genes in NTHi.

The lipooligosaccharide (LOS) of *Haemophilus influenzae* is an important factor in pathogenesis and virulence (10, 11, 36). The lipid A portion of LOS is responsible for the toxicity associated with this organism, but the role of the oligosaccharide portion of the LOS in pathogenesis of *H. influenzae* infection is less clear (10). The LOS of *H. influenzae* is analogous to the lipopolysaccharide (LPS) of enteric gram-negative bacteria; it contains lipid A linked by 3-deoxy-D-manno-octulosonic acid (KDO) to a heterogeneous sugar polymer, but the LOS lacks the repeating O-antigen units (Fig. 1) (6). The physicochemical analysis of the structure of oligosaccharide from *H. influenzae* LOS has shown a unique deep core structure with variable branched chains which makes it more complex than LOS from *Neisseria* and *Bordetella* species (6, 18).

The ability to study the genes required for LOS biosynthesis has been hampered by the lack of genetic tools for *Haemophilus* species. Several attempts have been made to elucidate the LOS biosynthesis genes (13, 32, 33). Recently, in our laboratory the *lsg* (LOS synthesis genes) cluster was isolated from *H. influenzae* type b (Hib) by using monoclonal antibodies to Hib LOS epitopes (1, 29). The *lsg* locus was shown to be able to modify the LPS of an *Escherichia coli* K-12 strain, and these modified LPS species were recognized by the monoclonal antibodies to the Hib LOS (29). It was also demonstrated by using isogenic Hib mutants that the *lsg* locus is involved in the Hib LOS biosynthesis (15). Sequence analysis of the 7.4-kb *lsg* locus revealed that it contains seven open reading frames (ORFs) and one partial ORF (14). The exact functions of these ORFs, however, have not yet been established.

The *rfa* genes, which encode LPS core biosynthesis enzymes, are present as a cluster on *Salmonella typhimurium* and *E. coli* K-12 chromosomes, and several *rfa* genes have been cloned and characterized. The *rfaE* gene of *S. typhimurium* is known to be located at 76 min on the genetic map outside of the *rfa* gene cluster (23). The *rfaE* mutant synthesizes heptose-deficient LPS (i.e., its LPS consists of only lipid A and KDO), and the *rfaE* gene is believed to be involved in the formation of ADP-heptose (27, 34), but cloning of the *rfaE* gene has not been reported yet.

The genetics of *Salmonella* LPS biosynthesis has been studied extensively during last two decades (25), and a great number of LPS mutant strains have been isolated and characterized (16, 21, 30, 34). The LPS of *Salmonella* species and the LOS of *Haemophilus* species share common structures in the deep core region, which makes it possible to isolate genes responsible for biosynthesis of this region from *Haemophilus* species by complementing *Salmonella* mutants.

To isolate genes involved in LOS biosynthesis of *H. influenzae*, we undertook a series of complementation studies in *S. typhimurium* LPS mutants with defined enzymatic defects. In this paper, we report the cloning, sequencing, and characterization of the nontypeable *H. influenzae* (NTHi) 2019 *rfaE* gene.

MATERIALS AND METHODS

Strains, plasmids, and culture conditions. The bacterial strains, phages, and plasmids used in this study are described in Table 1. *E. coli* strains were grown in LB medium, and *S. typhimurium* strains were grown in heart infusion (HI) or LB medium containing appropriate antibiotics at 37°C. NTHi 2019 was cultured in HI broth supplemented with 4% Fildes reagent (Difco Laboratories) (sHI) at 37°C with agitation or on sHI agar plates at 37°C in 5% CO_2 atmosphere. The

* Corresponding author. Mailing address: Department of Microbiology, University of Iowa, Bowen Science Building, 3-401, 51 Newton Road, Iowa City, IA 52242. Phone: (319) 335-7811. Fax: (319) 335-9006.

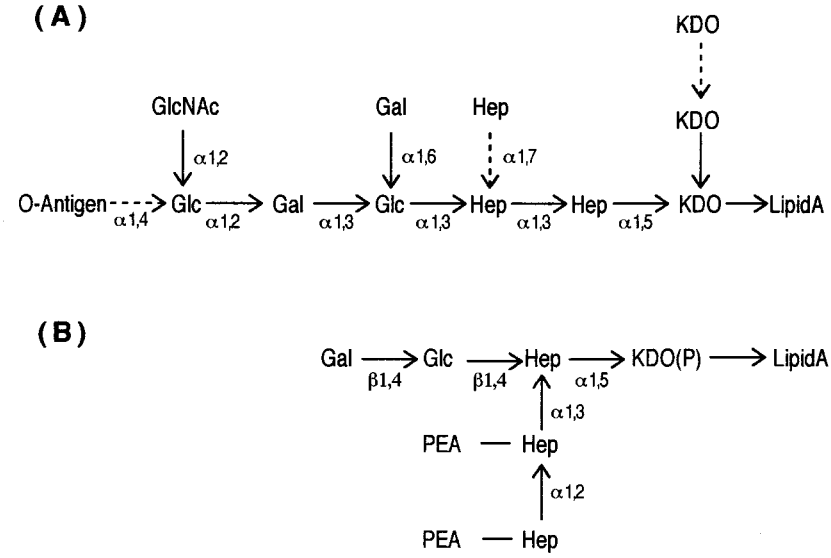

FIG. 1. Schematic illustration of the structures of *S. typhimurium* LPS (A) and NTHi 2019 LOS (B). Abbreviations: Gal, galactose; Glc, glucose; GlcNAc, *N*-acetylglucosamine; Hep, L-glycero-D-mannoheptose; KDO, 3-deoxy-D-manno-octulosonic acid; P, phosphate; PEA, phosphoethanolamine. Possible partial substitutions are indicated with dashed arrows. The higher-molecular-weight forms of NTHi 2019 LOS have not been fully characterized.

test for sensitivity to LPS-specific phages was performed by applying 5 μl of each phage stock suspension to lawns of *S. typhimurium* strains.

DNA manipulations. Restriction enzymes were purchased from New England Biolabs, and T4 DNA ligase was purchased from Promega. Standard DNA recombinant procedures were performed as described previously (22). Transformation of *E. coli* strains with plasmid DNA was routinely done by the $CaCl_2$ method (7); *Salmonella* strains were transformed by electroporation (2). For cloning of DNA fragments made by PCR into a vector DNA, the Promega pGEM-T vector system was used.

Construction of a genomic library. The chromosomal DNA from NTHi 2019 was partially digested with *Apo*I and resolved on a 0.8% agarose gel. A DNA fraction ranging from 4 to 8 kb was excised, purified by using DEAE-cellulose

TABLE 1. Bacterial and phage strains and plasmids used in this study

Strain or plasmid	Genotype or characteristic[a]	Source or reference
H. influenzae 2019	Wild type	M. A. Apicella
S. typhimurium LT2		
SL3770	*rfa*$^+$	SGSC[b]
SL1102	*rfaE543 metA22 trpC2 H1-b H2-e,n,x fla-66 rpsL120 xyl-404 metE551*	SGSC
SL3019	*rfaE827 rfaL446 SD14* (*E1*) *azi gal rha his*	SGSC
SA1377	*rfaC630*(P22)$^+$	SGSC
SL3600	*rfaD657 metA22 trpC2 H1-b H2-e,n,x fla-66 rpsL120 xyl-404 metE551*	SGSC
E. coli K-12		
DH5αF′	F′/*endA1 hsdR17* ($r_K^-m_K^+$) *supE44 thi-1 recA1 gyrA* (Nal^r) *relA1* Δ(*lacZYA-argF*)*U169 deoR* [ϕ80*dlac*Δ(*lacZ*)*M15*]	Gibco BRL
ER1647	F$^-$ λ$^-$ *trp-31 his-1 rpsL104* (Str^r) *fhuA2* Δ(*lacZ*)*r1 supE44 xyl-7 mtl-2 metB1 recD1014 mcrA1272*::Tn*10* Δ(*mcrB hsdRMS mrr*)2:Tn*10* Tet^r Str^r	Novagen
BM25.8	F′ *traD36 lacI*^{q} *lacZ*Δ*M15 proAB*/*supE thi* Δ(*lac-proAB*) λ*imm*^{434}(P1) Cm^r Kan^r	Novagen
Phage		
λSH*lox*	Cloning vector, Amp^r	Novagen
P22.c2	Smooth-specific phage of *S. typhimurium*	SGSC
Felix-O	Smooth-specific phage of *S. typhimurium*	SGSC
Ffm	Rough-specific phage of *S. typhimurium*	SGSC
Br60	Rough-specific phage of *S. typhimurium*	SGSC
Plasmids		
pUC19	Cloning and expression vector, Amp^r	Gibco BRL
pSH*lox*-1	Cloning vector, Amp^r	Novagen
pGEM-T	PCR cloning vector, Amp^r	Promega
pGEM-7Z(f+)	Cloning vector, Amp^r	Promega
pHSS6	Cloning vector, Kan^r	26
pHIE0	4-kb DNA from NTHi 2019 in pSH*lox*-1, Amp^r	This study
pHIE62	3.5-kb DNA from NTHi 2019 in pSH*lox*-1, Amp^r	This study
pHIE2C	Mini-Tn*3* inserted into *rfaE* open reading frame of 2.4-kb *Bgl*II fragment in pHSS6, Kan^r, Cm^r	This study
pHIE2K	Kanamycin resistance gene cassette cloned into *Sca*I site of pHIE2, Amp^r, Kan^r	This study

[a] Abbreviations for antibiotics: Amp, ampicillin; Cm, chloramphenicol; Kan, kanamycin; Str, streptomycin; Tet, tetracycline.
[b] SGSC, Salmonella Genetic Stock Center, University of Calgary, Calgary, Alberta, Canada.

membranes (Schleicher & Schuell), and ligated to *Eco*RI-digested λSH*lox* vector arms (Novagen, Inc.). The ligated mixture was packaged by using Gigapack II Gold packaging extract (Stratagene) according to the manufacturer's instructions.

The plasmid library was made by in vivo excision from the phage library as follows. The phage library stock containing approximately 8×10^5 PFU was mixed with 2 ml of the host strain BM25.8 (optical density at 600 nm of 0.5), which had been grown in LB with 10 mM $MgSO_4$, 0.2% maltose, 34 μg of chloramphenicol per ml, and 50 μg of kanamycin per ml, and incubated at 37°C for 30 min. One hundred microliters of the host-phage mixture was spread on a 150-mm-diameter LB agar plate containing 100 μg of carbenicillin per ml and allowed to grow overnight at 37°C. The colonies formed were pooled, and the plasmids were purified by the alkaline lysis method followed by CsCl gradient centrifugation (22).

Complementation of *S. typhimurium rfa* mutants. Overnight cultures of the *S. typhimurium rfa* mutants were inoculated into 100 ml of fresh HI medium and grown at 37°C with vigorous shaking to an optical density at 600 nm of 0.5. The cells were chilled on ice and centrifuged. The pellets were washed twice with ice-cold water and resuspended with a volume of 10% glycerol–water (vol/vol) equal to that of the pellet.

Twenty five microliters of the cells were electroporated with 10 ng of the plasmid library DNA by using a Cell Porater (Gibco BRL), incubated in 1 ml of HI medium at 37°C for 4 h with shaking, and then spread on HI agar plates containing 50 μg of ampicillin and 100 μg of novobiocin per ml. Plasmid DNA was purified from each transformant and retransformed to the mutant strain to confirm the complementation. The transformants were also tested for phage sensitivities and LPS phenotype.

DNA sequencing and analysis. DNA sequence was determined by the dideoxy-chain termination method (24), using a Sequenase version 2.0 sequencing kit (U.S. Biochemical) with primers commercially available or synthesized with a model 381A DNA synthesizer (Applied Biosystems, Inc.). DNA sequence was analyzed with the University of Wisconsin Genetics Computer Group package and GeneWorks software (IntelliGenetics, Inc.). Protein sequence alignment was done by using the BESTFIT program of the Genetics Computer Group sequence analysis software.

Genomic Southern hybridization. NTHi 2019 genomic DNA was digested with restriction enzymes, resolved on a 1% agarose gel, transferred to Hybond-N membrane (Amersham) by capillary blotting overnight, and cross-linked to the membrane by using a Stratalinker (Stratagene). After prehybridization, the membrane was hybridized with a digoxigenin-dUTP-labeled DNA probe and washed. The hybridized probe was detected by using a DIG Luminescent Detection Kit (Boehringer Mannheim Biochemicals), and the membrane was exposed to Kodak X-Omat AR film at room temperature.

LPS gel analysis. LPS of *S. typhimurium* was prepared by extraction with phenol and precipitation in ethanol as described previously (5), separated on a 14% polyacrylamide gel containing sodium dodecyl sulfate (SDS) (12), and visualized by silver staining as described previously (31).

Primer extension analysis. RNA was extracted from NTHi 2019 grown to an optical density at 600 nm of 0.6 in sHI broth by sonication, digestion with proteinase K, phenol-chloroform extraction, and ethanol precipitation as described elsewhere (2). The purified RNA was quantitated spectrophotometrically, and the quality of RNA was confirmed on a formaldehyde-agarose gel stained with ethidium bromide.

Primer extension analysis was carried out by using a Promega primer extension kit as instructed by the manufacturer except for annealing, which was done by heating the reaction mixture to 70°C, incubating it at 60°C for 20 min, and then slowly cooling it to room temperature. Twenty micrograms of RNA was used for each reaction, and the reaction products were precipitated in ethanol, dissolved in loading dye, and loaded on a 6% sequencing gel. The dideoxy sequencing ladder with the same template and the same primer was used as a marker to confirm the position of the primer-extended products.

In vitro transcription-translation analysis. Plasmids were purified by CsCl gradient centrifugation and used for in vitro transcription-translation analysis (Promega). The translation products were labeled with translation-grade [^{35}S]methionine (Amersham). The reaction mixtures were separated on an SDS–11% polyacrylamide gel (12), which was dried and exposed to X-ray film. ^{14}C-labeled protein molecular weight Rainbow standards (Amersham) were used to generate a linear regression curve between 14.3 and 200 kDa for determination of relative molecular weights.

Nucleotide sequence accession number. The DNA sequence of the *rfaE* gene and the adjacent region has been submitted to GenBank and assigned accession number U17642.

RESULTS

Isolation of a plasmid carrying the *rfaE* gene from an NTHi 2019 library. A plasmid carrying the *rfaE* gene was isolated from an NTHi 2019 library by complementing the *rfaE* mutant of *S. typhimurium* LT2. To accomplish this, the NTHi 2019 plasmid library DNA was transformed into strain SL1102, which is *rfaE* and thus defective in ADP-heptose synthesis (27, 34), resulting in an incomplete LPS core. Transformants carrying a plasmid containing the *rfaE* gene were selected on the basis of the properties of *Salmonella* strains with a wild-type LPS that are less permeable and thus more resistant to hydrophobic antibiotics than mutant strains with a defective LPS core structure (21). Cells were plated on an HI plate containing ampicillin (50 μg/ml) and a hydrophobic antibiotic, novobiocin (100 μg/ml). Colonies that grew on this plate were retested for resistance to novobiocin by growth in liquid culture in the presence of novobiocin. These transformants were also tested for sensitivity to the LPS-specific phages. All of them were resistant to the rough-specific phages, Br60 and Ffm, and sensitive to the smooth-specific phages, Felix-O and P22.c2 (Table 2), indicating that they synthesized a complete LPS core structure and O-antigen repeating units. Plasmid DNA was extracted from one of the clones and transformed into SL1102 with selection for ampicillin resistance. All of the transformants tested grew in the presence of novobiocin, indicating that the plasmid carried a gene conferring resistance to novobiocin. Restriction mapping revealed that this plasmid contained an insert of 4 kb, which was designated pHIE0 (Fig. 2). LPS gel analysis confirmed that the LPS of SL1102 complemented with plasmid pHIE0 was converted to the wild-type phenotype (see Fig. 8).

TABLE 2. Phage sensitivity of *S. typhimurium rfa* mutants complemented with various plasmids carrying NTHi 2019 DNA

Strain	Genotype	Plasmid	Phage sensitivity[a]			
			P22.c2	Felix-O	Ffm	Br60
SL3770	*rfa*$^+$		+	+	−	−
SL1102	*rfaE543*		−	−	+	+
SL1102	*rfaE543*	pHIE0	+	+	−	−
SL1102	*rfaE543*	pHIE1	−	−	+	+
SL1102	*rfaE543*	pHIE2	+	+	−	−
SL1102	*rfaE543*	pHIE3	+	+	−	−
SL1102	*rfaE543*	pHIE4	−	−	+	+
SL1102	*rfaE543*	pHIE5	−	−	+	+
SL1102	*rfaE543*	pHIE6	−	−	+	+
SL1102	*rfaE543*	pHIE7	+	+	−	−
SL1102	*rfaE543*	pHIE2C	−	−	+	+
SL3019	*rfaE827rfaL446*		−	−	+	+
SL3019	*rfaE827rfaL446*	pHIE62	+	+	+	+
SL3019	*rfaE827rfaL446*	pHIE0	+	+	+	+
SA1377	*rfaC630*		−	−	+	+
SA1377	*rfaC630*	pHIE0	−	−	+	+
SL3600	*rfaD657*		−	−	+	+
SL3600	*rfaD657*	pHIE0	−	−	+	+

[a] +, sensitive; −, resistant. Phage P22.c2 requires O antigen, Felix-O requires a complete core, Ffm requires a complete lack of O antigen for sensitivity, and Br60 recognizes inner core structure.

Another *rfaE* mutant strain, SL3019, was also used to isolate the *rfaE* gene from NTHi 2019 by the procedure described above. One of the clones isolated contained a 3.5-kb NTHi 2019 DNA fragment and was designated pHIE62. SL3019 transformed with pHIE62 became sensitive to phages P22.c2 and Felix-O but also retained the sensitivity to Ffm and Br60. SL3019 complemented with pHIE0 showed the same pattern of phage sensitivity as seen with pHIE62 (Table 2).

Two other *Salmonella* LPS mutant strains, SA1377 and SL3600, have the same LPS phenotype as SL1102, i.e., make heptoseless LPS, but have mutations in other genes, *rfaC* and *rfaD*, encoding ADP-heptose:lipopolysaccharide heptosyltransferase I and ADP-L-glycero-D-mannoheptose-6-epimerase, respectively (4, 17, 27). When these strains were trans-

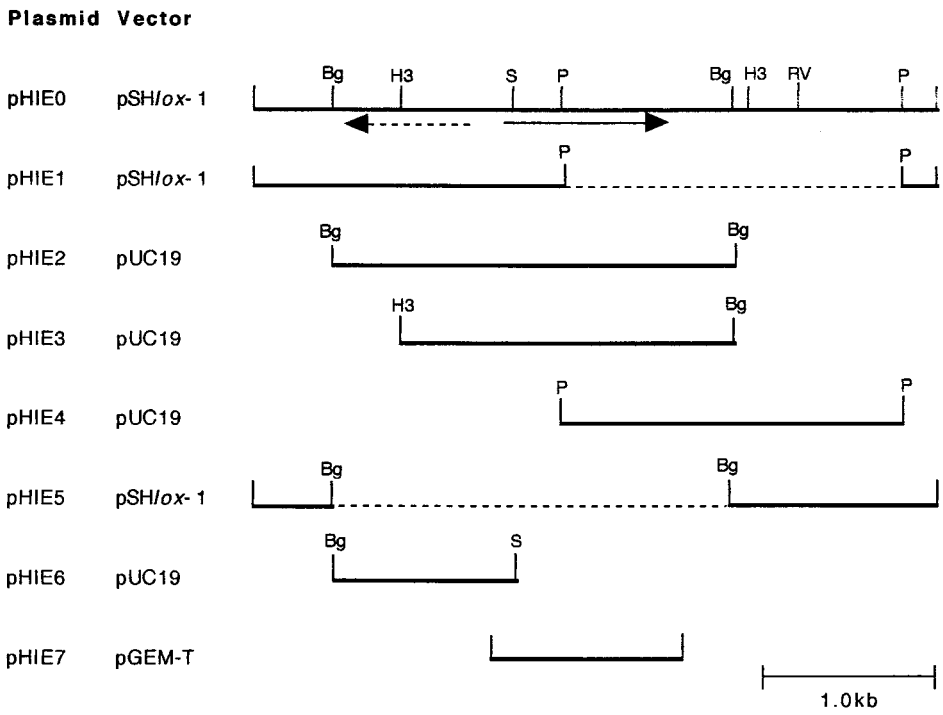

FIG. 2. Restriction map of the subclones of the NTHi 2019 *rfaE* clone. The solid lines indicate the NTHi 2019 DNA, and the dashed lines indicate deleted sequences. The solid arrow indicates *rfaE* ORF, and the dashed arrow indicates the *htrB* ORF. Only restriction sites relevant to the plasmid constructions are shown. Bg, *Bgl*II; H3, *Hin*dIII; RV, *Eco*RV; P, *Pvu*II; S, *Sca*I.

formed with pHIE0, neither of them was complemented, as determined by phage sensitivity (Table 2).

Subcloning of the plasmid carrying the *rfaE* gene. To delimit the *rfaE* gene region, a series of subclones was made from pHIE0 (Fig. 2). Plasmids pHIE1 and pHIE5 were made by deleting 2.0-kb *Pvu*II and 2.4-kb *Bgl*II fragments from pHIE0, respectively. pHIE2, pHIE3, pHIE4, and pHIE6 were generated by subcloning 2.4-kb *Bgl*II, 2.0-kb *Hin*dIII-*Bgl*II, 2.0-kb *Pvu*II, and 1.2-kb *Bgl*II-*Sca*I fragments into the vector pUC19, respectively. These subclones were tested for the ability to complement SL1102. SL1102 was transformed with each subclone and tested for resistance to novobiocin and changes in the sensitivities to the LPS-specific phages (Table 2). pHIE2 and pHIE3 were able to complement SL1102.

To confirm that the *rfaE* gene clone isolated is genuine NTHi 2019 genomic DNA, genomic Southern hybridization was carried out as described in Materials and Methods. The probe used for hybridization was a PCR fragment made with primers C and D (the sequences of which are shown in Table 3). As seen in Fig. 3, the *Hin*dIII digest yielded a 2.0-kb fragment, the *Bgl*II digest yielded a 2.4-kb fragment, and the *Pvu*II digest yielded a 2.0-kb fragment, while *Bam*HI and *Eco*RV digests produced single bands hybridizing with the probe, which was consistent with the results expected from the restriction pattern of pHIE0.

TABLE 3. Sequences of the primers used in this study

Primer	Sequence
A	5′-TGGGGCTAAAAAGTGCGGTTGAAATTGAGG-3′
B	5′-AAGCATCACATCGCCTAATACAA-3′
C	5′-CCCATCCAACCATTACTAAA-3′
D	5′-AAATATCGAAACAGCCATTAG-3′

Sequence analysis of the *rfaE* gene and determination of the promoter region. Sequence analysis of the 2.4-kb *Bgl*II fragment revealed two ORFs, one of which encoded a polypeptide of 311 amino acid residues highly homologous to the *E. coli* HtrB protein (56% identity and 74% similarity) (9). The other ORF encoded a polypeptide of 342 amino acid residues. A probable Shine-Dalgarno sequence was identified at 8 bp upstream of the predicted translation initiation codon ATG of

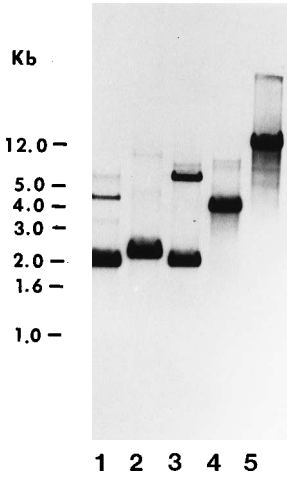

FIG. 3. Genomic Southern hybridization of NTHi 2019. Genomic DNAs were digested with various restriction enzymes, resolved on a 1% agarose gel, blotted to a nylon membrane, and hybridized with the 2.4-kb *Bgl*II fragment after labeling with digoxigenin-dUTP by random priming. The sizes of DNA markers are indicated on the left. Lanes: 1, *Hin*dIII; 2, *Bgl*II; 3, *Pvu*II; 4, *Eco*RV; 5, *Bam*HI.

```
HiRFAE      KQAKVLVLGDVMLDRYWFGATNRISPERPVPVVRVQENEERAGGAANVAM 58
            ... |:::|.. :|   :: ..   ::. .. ::.    ..:|.: |:|
EcFRU       EKDYVVIIGSANIDVAGYSHESLNYADSNPGKIKF....TPGGVGRNIAQ 48

HiRFAE      NIASLNVPVQLMGLIGQDETG.SALSHLLEKQKIDCNFVALETHPTITKL 107
            |:| |. .. |:: :| |  | | |.:  :.. .  . : :... | . |
EcFRU       NLALLGNKAWLLSAVGSDFYGQSLLTQTNQSGVYVDKCLIVPGENTSSYL 98

HiRFAE      RILSRHQQLLRLDFEEDFNNVDCKDLLAKLESAVKNYGALILSDYGKGTL 157
            .:|.. .::|    : ::.|. . :.||. :. :.. : :|:.|:. :.
EcFRU       SLLDNTGEMLVAINDMNISNAITAEYLAQHGEFIQR.AKVIVADCNISE. 146

HiRFAE      KDVQKMIQIARKGNVPVLIDPKGT.....DFERYRGATLLTPNMSEFEAV 202
            ..:..::: |  :|||::|| :.      :|...  |.||. | |.:
EcFRU       EALAWILDNA..ANVPVFVDPVSAWKCVKVRDRLNQIHTLKPNRLEAETL 194

HiRFAE      VGKCNTEEEIIKKGLKLISDIELTALLVTRSEKGMTLLRPNQEPYHLPTV 252
              . .: | : |.  :: : :|. |::..::.|: .   ..|.   :.:
EcFRU       SGIALSGREDVAKVAAWFHQHGLNRLVLSMGGDGVYYSDISGESGWSAPI 244

HiRFAE      AKEVFDVTGAGDTVISVLATTLADGCSFEESCYLANVAAGIVVGKLGTST 302
             .:|::|||||.::..||..:.|| .|.||. :|....::.::
EcFRU       KTNVINVTGAGDAMMAGLASCWVDGMPFAESVRFAQGCSSMALS..... 288

HiRFAE      VSTVELENAIHARPETGFGIMSEAE 327
               .|..|. . . ...::::.:||
EcFRU       ...CEYTNNPDLSIANVISLVENAE 310
```

FIG. 4. Comparison of the predicted amino acid sequences of the *rfaE* gene product of NTHi 2019 (HiRFAE) and the ORF present downstream of *E. coli* fructose phosphotransferase system genes (EcFRU). Bars indicate identical amino acid residues, and dots indicate similar residues. Parts of the N- and C-terminal sequences are not shown because of lack of homology.

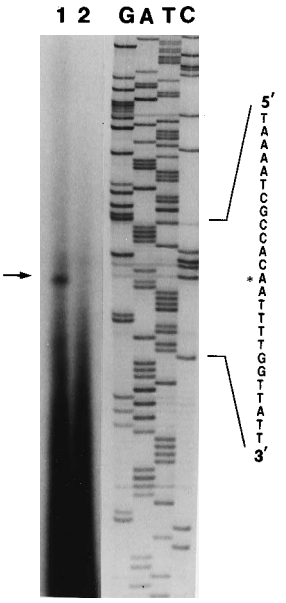

FIG. 5. Primer extension analysis of the *rfaE* gene of NTHi 2019. Total RNA was isolated from the log-phase culture of NTHi 2019 grown in sHI medium at 37°C. The sequence of the primer used (primer B) is shown in Table 3. The position of the transcription start site is indicated with an asterisk on the sequence, and the position for the extended product is indicated with an arrow. The DNA sequencing ladders (GATC) were obtained by using the same primer as that in the primer extension reaction. Lanes: 1, 20 μg of RNA; 2, no RNA.

this ORF. To construct a plasmid containing this putative *rfaE* gene ORF, PCR was carried out with two oligonucleotides (primers A and D), one upstream of the promoter region and the other downstream of the translation stop codon. The resulting 1.3-kb PCR fragment was cloned into pGEM-T vector, yielding pHIE7. SL1102 complemented with pHIE7 was shown to become sensitive to the smooth-specific phages, Felix-O and P22.c2 (Table 2). Conversion of the LPS phenotype of this strain to wild type was also confirmed by LPS gel analysis. Disruption of this ORF in pHIE2C by insertion of mini-Tn*3* abolished the complementing ability of the plasmid (Table 2), which indicated that this ORF was the authentic *rfaE* gene. Plasmid pHIE62 was partially sequenced by using an oligonucleotide (primer C) located in the ORF of the *rfaE* gene from pHIE0. The nucleotide sequence over 300 bp was 100% homologous to the sequence of the *rfaE* gene from pHIE0, suggesting that these two clones contained identical genes. A database search for sequence homology revealed that an ORF present downstream of the *E. coli* fructose phosphotransferase system genes exhibited homology with the product of the *rfaE* gene (21% identity and 45% similarity) (19). Amino acid sequence comparison of these two genes is shown in Fig. 4.

Primer extension analysis was carried out to determine the promoter region of the *rfaE* gene (Fig. 5). A single transcription start site was found, which was designated +1 in Fig. 6. The first nucleotide transcribed was located at 32 bp upstream of the predicted translation start site. The region upstream of the transcription start site contained a sequence (TAAAAT) homologous to the −10 region of the bacterial σ^{70}-dependent promoters at an appropriate distance (7 bp), but no sequence resembling the consensus sequence of the −35 region (TTGACA) was found (Fig. 6).

Identification of the *rfaE* gene product. A coupled transcription-translation system was used to confirm the production of a protein from the *rfaE* gene (Fig. 7). SDS-polyacrylamide gel electrophoresis (PAGE) analysis showed that three proteins were produced from the 4-kb NTHi 2019 DNA (lane 3). The deduced molecular mass of the RfaE protein was 37.6 kDa. However, two bands produced from pHIE7 had molecular masses of approximately 37 and 39 kDa on a gel (lane 2). Considering that the β-lactamase precursor protein ran with an apparent molecular mass of 33 kDa, higher than the actual molecular mass of 31.5 kDa on the gel, however, this would be the position of a protein of 37.6 kDa deduced from the nucleotide sequence of the *rfaE* gene. Translation using the internal ATG codon would give a protein of 35.5 kDa, causing the doublet with pHIE7.

LPS phenotypes of the *rfaE* mutant complemented with the plasmid carrying the NTHi 2019 *rfaE* gene. LPS from the *rfaE* mutant carrying the NTHi 2019 *rfaE* gene was analyzed by SDS-PAGE followed by silver staining. LPS from SL3770, which is *rfa*$^+$, showed a wild-type phenotype (Fig. 8, lane 1). LPS from SL1102 contained rapidly migrating bands representing heptose-deficient incomplete core structures (lane 2). SL1102 complemented with pHIE0 showed LPS which was indistinguishable from the wild-type LPS (lane 3). SL1102 complemented with pHIE7 showed a pattern similar to that of SL1102 carrying pHIE0, but it still contained the band corre-

```
GGGGCGTAGTTTATCGTAAATTTGAGTTGTGGTAAAATCGCCACAATTTT
                                  -10        +1

GGTTATTCAAATAGAGAGATTTTAAAATGGCTCAATATTCAGCAGAATTT
               SD          M  A  Q  Y  S  A  E  F

AAGCAAGCAAAAGTACTTGTATTAGGCGATGTGATGCTTGATCGTTATTG
K  Q  A  K  V  L  V  L  G  D  V  M  L  D  R  Y  W
```

FIG. 6. The promoter region of the *rfaE* gene. The arrow indicates the position of the transcription start site determined by primer extension analysis as shown in Fig. 5, and the −10 region for the promoter is double underlined. The sequence for the primer used for primer extension analysis is underlined. The probable Shine-Dalgarno ribosome-binding site (SD) is identified upstream of the translation initiation codon ATG.

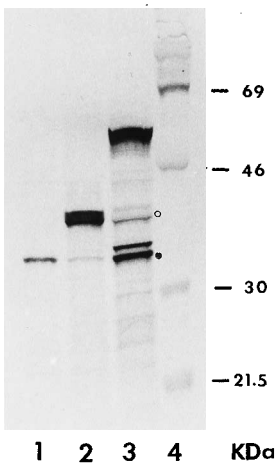

FIG. 7. In vitro transcription and translation of the *rfaE* gene. Lanes: 1, pGEM7Zf(+); 2, pHIE7; 3, pHIE0; 4, protein molecular weight standards. The open circle indicates protein products corresponding to the *rfaE* gene; the closed circle indicates the β-lactamase precursor.

sponding to the low-molecular-weight incomplete core structure (lane 4), suggesting that the level of the *rfaE* gene expression is lower from pHIE7.

LPS from SL3019 complemented with pHIE62 showed two bands, one corresponding to the incomplete core LPS and the other corresponding to the complete core LPS, but no slowly migrating ladders were found (data not shown). Transformation of SL3019 with pHIE0 produced the same result as transformation with pHIE62.

DISCUSSION

The structural similarities of the enterobacterial LPS and the *Haemophilus* LOS have allowed us to isolate the *Haemophilus rfaE* gene from an NTHi 2019 library by using the *S. typhimurium rfaE* mutant. From the results of novobiocin sensitivity, phage typing, and LPS gel analysis, we conclude that the isolated *Haemophilus* gene is able to complement the *rfaE* mutation of *S. typhimurium* and is thus an *rfaE* homolog of NTHi 2019. Other LPS mutants, the *rfaC* and the *rfaD* mutant strains, which have the same LPS phenotype but whose mutations lie in other genes failed to show complementation with plasmid pHIE0 (Table 2). This experiment indicated that the NTHi 2019 clone is not *rfaC* or *rfaD*.

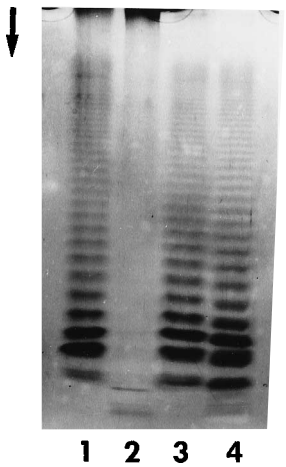

FIG. 8. Silver-stained LPS gel showing complementation of the *S. typhimurium rfaE* mutant by plasmids carrying the *H. influenzae rfaE* gene. The arrow indicates the direction of sample migration. Lanes: 1, SL3770 (*rfa*$^+$); 2, SL1102 (*rfaE543*); 3, SL1102 transformed with plasmid pHIE0; 4, SL1102 transformed with plasmid pHIE7. Note that lane 4, but not lane 3, contains the low-molecular-weight LPS band as seen in lane 2.

Though both plasmids pHIE2 and pHIE3, carrying the 2.4-kb *Bgl*II fragment and the 2.0-kb *Hin*dIII-*Bgl*II fragment from pHIE0, respectively, were able to complement the *S. typhimurium rfaE* phenotype, SL1102 transformed with pHIE3 grew more slowly in broth in the presence of novobiocin than SL1102 transformed with pHIE2 or SL3770 (data not shown). pHIE7 containing the 1.3-kb fragment was sufficient for the complementation (Table 2). These data suggested that the region upstream of the *rfaE* gene contains the sequence which may be important to novobiocin resistance. Since pHIE6 containing the upstream *Bgl*II-*Sca*I fragment did not show any ability to complement the *rfaE* phenotype or to give resistance to novobiocin, the effect of this region on the novobiocin sensitivity of the *rfaE* mutant is thought to be indirect. The gel pattern of LPS from pHIE7 suggested that *rfaE* gene expression is lower from pHIE7 than from pHIE0 (Fig. 8). Deletion of the upstream region of the *rfaE* gene which contains the *htrB* homolog may be responsible for the lower expression. The *E. coli* HtrB protein was first identified to be essential for cell viability at high temperature and also to affect the LPS silver-staining pattern on SDS-PAGE and thus was suggested to be involved in LPS biosynthesis (8, 9, 25). It is interesting that in NTHi 2019 *rfaE* and *htrB* share the promoter region and are transcribed in opposite directions, suggesting that the mechanisms of regulation of their expression are related.

Another *rfa* gene responsible for synthesis of heptose is the *rfaD* gene, which encodes the enzyme ADP-L-glycero-D-mannoheptose-6-epimerase (17). The *rfaD* gene has been cloned and characterized in both *E. coli* K-12 and *S. typhimurium* (20, 28). Their sequences are highly homologous to each other, but their gene expression seems to be regulated in different ways. The *E. coli rfaD* gene was identified in the process of screening for heat shock proteins and designated *htrM* (20). Its upstream region contains three promoters, one of which is a σ^{70}-dependent promoter and one of which is a σ^{32}-dependent heat shock promoter, while the *S. typhimurium rfaD* gene lacks the heat shock promoter (28). The *rfaE* gene is also involved in heptose synthesis, but primer extension analysis showed that the NTHi 2019 *rfaE* gene had only a σ^{70}-dependent promoter, and its expression was not affected by temperature (data not shown).

The database search for the *rfaE* homolog revealed one ORF encoding a 33.6-kDa protein in proximity to the phosphoenolpyruvate-dependent sugar phosphotransferase system gene cluster of *E. coli*. This ORF is present at the 3′ end of the *fruA* gene, which encodes a fructose phosphotransferase. However, there is a transcription termination sequence downstream of the *fruA* gene, and it is thought that the ORF belongs to a different operon and may not be a part of the fructose transport or fructose phosphotransferase system genes (19). Currently, the function of this ORF is not known. Considering its location and the homology to the NTHi 2019 *rfaE* gene, however, it is tempting to speculate that *fruA* encodes a protein with a function similar to that of the RfaE protein, i.e., phosphorylation of sugars.

These experiments demonstrated that complementation of the defined LPS mutants of *S. typhimurium* can be used to isolate LOS synthesis genes from *Haemophilus* species. It is thought that the NTHi 2019 *rfaE* gene complementing the *S. typhimurium rfaE* mutant performs a similar function in NTHi

2019. To determine whether the *rfaE* gene is directly involved in LOS biosynthesis in *H. influenzae*, we performed a mutagenesis study with strains NTHi 2019 and Hib A2. pHIE2C, constructed by transposon mutagenesis using mini-Tn*3* (26), and pHIE2K, constructed by insertion of a kanamycin resistance gene cassette into the *rfaE* gene (Fig. 2), were used to transform *H. influenzae*. Since it is known that *S. typhimurium* strains with truncated LPS show temperature sensitivity (3), we attempted to rescue mutants by incubating cells at 30°C. We were unsuccessful in recovering an *rfaE* mutant of strains 2019 and Hib A2. We could rescue an *rfaD* mutant of strain 2019 which has the same LOS phenotype (KDO-lipid A) as would be expected with an ADP-heptose synthase mutant obtained by using this approach (unpublished data). This finding suggests that accumulation of substrate(s) in the *rfaE* mutant is responsible for the lethal effect of the mutation rather than modification of the cell wall due to the LOS truncation.

While a similar approach using complementation of *S. typhimurium* mutants and conversion to a novobiocin-resistant phenotype has been used for *E. coli* (4), this technique for selection of an LPS phenotype has only recently been applied to nonenterobacterial species (35). The difficulties inherent in the molecular manipulation of organisms such as *Neisseria* and *Haemophilus* species make this approach particularly attractive in the identification and cloning of LOS genes.

ACKNOWLEDGMENT

This work was supported by Public Health Service grant AI 24616.

REFERENCES

1. **Abu Kwaik, Y., R. E. McLaughlin, M. A. Apicella, and S. M. Spinola.** 1991. Analysis of *Haemophilus influenzae* type b lipooligosaccharide-synthesis genes that assemble or expose a 2-keto-3-deoxyoctulosonic acid epitope. Mol. Microbiol. **5:**2475–2480.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. John Wiley & Sons, New York.
3. **Chatterjee, A. K., K. E. Sanderson, and H. Ross.** 1976. Influence of temperature on growth of lipopolysaccharide-deficient (rough) mutants of *Salmonella typhimurium* and *Salmonella minnesota*. Can. J. Microbiol. **22:**1540–1548.
4. **Coleman, W. G.** 1983. The *rfaD* gene codes for ADP-L-glycero-D-mannoheptose-6-epimerase. J. Biol. Chem. **258:**1985–1990.
5. **Galanos, C., O. Luderitz, E. T. Rietschel, and O. Westphal.** 1969. A new method for the extraction of R lipopolysaccharides. Eur. J. Biol. **9:**245–249.
6. **Gibson, B. W., W. Melaugh, N. J. Phillips, M. A. Apicella, A. A. Campagnari, and J. M. Griffiss.** 1993. Investigation of the structural heterogeneity of lipooligosaccharides from pathogenic *Haemophilus* and *Neisseria* species and of R-type lipopolysaccharides from *Salmonella typhimurium* by electrospray mass spectrometry. J. Bacteriol. **175:**2702–2712.
7. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
8. **Karow, M., O. Fayet, A. Cegielska, T. Ziegelhoffer, and C. Georgopoulos.** 1991. Isolation and characterization of the *Escherichia coli htrB* gene, whose product is essential for bacterial viability above 33°C in rich media. J. Bacteriol. **173:**741–750.
9. **Karow, M., and C. Georgopoulos.** 1991. Sequencing, mutational analysis, and transcriptional regulation of the *Escherichia coli htrB* gene. Mol. Microbiol. **5:**2285–2292.
10. **Kimura, A., and E. J. Hansen.** 1986. Antigenic and phenotypic variations of *Haemophilus influenzae* type b lipopolysaccharide and their relationship to virulence. Infect. Immun. **51:**69–79.
11. **Kimura, A., C. C. Patrick, E. E. Miller, L. D. Cope, G. H. McCracken, Jr., and E. J. Hansen.** 1987. *Haemophilus influenzae* type b lipopolysaccharide: stability of expression and association with virulence. Infect. Immun. **55:** 1979–1986.
12. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
13. **Maskell, D. J., M. J. Szabo, P. D. Butler, A. E. Williams, and E. R. Moxon.** 1991. Molecular analysis of a complex locus from *Haemophilus influenzae* involved in phase-variable lipopolysaccharide biosynthesis. Mol. Microbiol. **5:**1013–1022.
14. **McLaughlin, R., N.-G. Lee, Y. Abu Kwaik, S. M. Spinola, and M. A. Apicella.** 1994. Characterization and sequence analysis of the *lsg* (LOS synthesis genes) locus from *Haemophilus influenzae* type b. J. Endotoxin Res. **1:**165–174.
15. **McLaughlin, R., S. M. Spinola, and M. A. Apicella.** 1992. Generation of lipooligosaccharide mutants of *Haemophilus influenzae* type b. J. Bacteriol. **174:**6455–6459.
16. **Naide, Y., H. Nikaido, P. H. Makela, R. G. Wilkinson, and B. A. D. Stocker.** 1965. Semirough strains of Salmonella. Proc. Natl. Acad. Sci. USA **53:**147–153.
17. **Pegues, J. C., L. Chen, A. W. Gordon, L. Ding, and W. G. Coleman, Jr.** 1990. Cloning, expression and characterization of the *Escherichia coli* K-12 *rfaD* gene. J. Bacteriol. **172:**4652–4660.
18. **Phillips, N. J., M. A. Apicella, J. M. Griffiss, and B. W. Gibson.** 1993. Structural studies of the lipooligosaccharides from *Haemophilus influenzae* type b strain A2. Biochemistry **32:**2003–2012.
19. **Prior, T. I., and H. L. Kornberg.** 1988. Nucleotide sequence of *fruA*, the gene specifying enzyme II^{fru} of the phosphoenolpyruvate-dependent sugar phosphotransferase system in *Escherichia coli* K12. J. Gen. Microbiol. **134:**2757–2768.
20. **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.
21. **Roantree, R. J., T. Kuo, and D. G. MacPhee.** 1977. The effect of defined lipopolysaccharide core defects upon antibiotic resistances of *Salmonella typhimurium*. J. Gen. Microbiol. **103:**223–234.
22. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Sanderson, K. E., and J. R. Roth.** 1988. Linkage map of *Salmonella typhimurium*, edition VII. Microbiol. Rev. **52:**485–532.
24. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
25. **Schnaitman, C. A., and J. D. Klena.** 1993. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. Microbiol. Rev. **57:**655–682.
26. **Seifert, H. S., M. So, and F. Heffron.** 1986. Shuttle mutagenesis: a method of introducing transposons into transformable organisms, p. 123–133. *In* H. S. Seifert, M. So, and F. Heffrons (ed.), Genetic engineering, principles and methods, vol. 8. Plenum Press, New York.
27. **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.
28. **Sirisena, D. M., P. R. MacLachlan, S.-L. Liu, A. Hessel, and K. E. Sanderson.** 1994. Molecular analysis of the *rfaD* gene, for heptose synthesis, and the *rfaF* gene, for heptose transfer, in lipopolysaccharide synthesis in *Salmonella typhimurium*. J. Bacteriol. **176:**2379–2385.
29. **Spinola, S. M., Y. Abu Kwaik, A. J. Lesse, A. A. Campagnari, and M. A. Apicella.** 1990. Cloning and expression in *Escherichia coli* of a *Haemophilus influenzae* type b lipooligosaccharide synthesis gene(s) that encodes a 2-keto-3-deoxyoctulosonic acid epitope. Infect. Immun. **58:**1558–1564.
30. **Subbaiah, T. V., and B. A. D. Stocker.** 1964. Rough mutants of *Salmonella typhimurium*. (I) Genetics. Nature (London) **201:**1298–1299.
31. **Tsai, C.-M., and C. E. Frasch.** 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal. Biochem. **119:**115–119.
32. **Weiser, J. N., A. A. Lindberg, E. J. Manning, E. J. Hansen, and E. R. Moxon.** 1989. Identification of a chromosomal locus for expression of lipopolysaccharide epitopes in *Haemophilus influenzae*. Infect. Immun. **57:**3045–3052.
33. **Weiser, J. N., J. M. Love, and E. R. Moxon.** 1989. The molecular mechanism of phase variation of *H. influenzae* lipopolysaccharide. Cell **59:**657–665.
34. **Wilkinson, R. G., P. Gemski, 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.
35. **Zhou, D., N.-G. Lee, and M. A. Apicella.** 1994. Lipooligosaccharide biosynthesis in *Neisseria gonorrhoeae*: cloning, identification and characterization of the α1,5 heptosyltransferase gene (*rfaC*). Mol. Microbiol. **14:**609–618.
36. **Zwahlen, A., L. G. Rubin, and E. R. Moxon.** 1986. Contribution of lipopolysaccharide to pathogenicity of *Haemophilus influenzae*: comparative virulence of genetically-related strains in rats. Microb. Pathog. **1:**465–473.