Characterization of Repetitive Sequences Controlling Phase Variation of *Haemophilus influenzae* Lipopolysaccharide

JEFFREY N. WEISER,† DUNCAN J. MASKELL, PETER D. BUTLER, ALF A. LINDBERG,‡ AND E. RICHARD MOXON,*

Molecular Infectious Diseases Group, Department of Paediatrics, Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DU, United Kingdom

Received 18 January 1990/Accepted 27 March 1990

Phase variation of lipopolysaccharide epitopes of an *Haemophilus influenzae* serotype b strain (strain RM.7004) occurs through a mechanism which depends on multiple tandem repeats of the DNA sequence 5'-CAAT-3' situated within the chromosomal locus lic1. We report here that the same tetranucleotide repeats are also found in two other genomic loci (lic2 and lic3) of RM.7004. Similar to lic1, there are multiple tandem repeats of 5'-CAAT-3' present at the 5' ends of long open reading frames in lic2 and lic3. Variation in the number of repeats of CAAT, by shifting the upstream initiation codons in or out of phase with the remainder of the open reading frame, could switch on or off the translation of downstream genes. Similar to previously reported findings for lic1, site-directed mutations in the open reading frame downstream (3') from the repeats of CAAT in lic2 abolished phase variation and identified DNA sequences required for the expression of additional oligosaccharide epitopes. When we used an oligonucleotide comprising five repeats of CAAT or DNA sequences specific for lic1, lic2, and lic3 as probes, a survey of other encapsulated *H. influenzae* strains (serotypes b through f) and non-encapsulated *H. influenzae* strains (including biotype aegyptius) showed that the chromosome of *H. influenzae* can have from two to five regions which contain multiple tandem repeats of CAAT in addition to other sequences which hybridize to lic1 and lic2.

*Haemophilus influenzae* lipopolysaccharide (LPS) is characterized by both inter- and intrastrain heterogeneity of the surface-exposed, neutral sugars of its core oligosaccharides (9, 12, 28). The intrastrain variation is due, at least in part, to a pattern of antigen switching (12) referred to as phase variation. Strains of *H. influenzae* undergo spontaneous, high-frequency gain and loss of reactivity with monoclonal antibodies (MAbs) which define epitopes of the core oligosaccharides of the LPS. Through phase variation of multiple epitopes the bacterium can express a diverse, but limited, number of different surface structures (28). Phase variants which express specific epitopes have enhanced virulence in an animal model (12) and are selected for or induced in the course of systemic infections of humans (29). The structure of one of these epitopes has been identified as Galla1-Gallb (M. Virji, J. N. Weiser, and E. R. Moxon, unpublished data). This disaccharide, which is also a phase-variable epitope on the LPS of *Neisseria gonorrhoeae*, is found on the glycosphingolipids of some human epithelial cells (4, 14). We have hypothesized that *H. influenzae* LPS is involved in an adaptive response which involves molecular mimicry of surface structures found on the epithelial cells of humans, its obligate host.

We have suggested a mechanism for the control of *H. influenzae* LPS phase variation. A chromosomal locus, lic1 (lipopolysaccharide core) (28, 29), encodes four proteins which mediate the biosynthesis and phase variation of oligosaccharide epitopes defined by MAbs. At the 5' end of the open reading frame of the first gene (licA) of the putative lic1 operon, there are multiple tandem repeats of the tetramer 5'-CAAT-3'. Variation in the number of copies of CAAT occurs spontaneously and involves a rec-independent mechanism, possibly slip-stranded mispairing (13). This shifts upstream initiation codons into or out of phase with the remainder of the open reading frame. Translation of licA, which is determined by the number of copies of CAAT, is required for phase variation of specific epitopes; expression of these core oligosaccharides, defined by MAbs 6A2 and 12D9, depends on downstream genes of the lic1 operon, which are designated licC and licD, respectively. Three levels of reactivity with these MAbs (+, +, +, +, +, and −) correlate with the number of CAAT repeats and the three different phases of translation (29). Since the number (n) of copies of CAAT in lic1 is variable, the repetitive region is designated (CAAT)n.

However, phase variation for at least five different LPS epitopes occurs in strain RM.7004 (28, 29), the serotype b strain from which lic1 was cloned; only two of these epitopes are determined by lic1. This suggests that other sequences may be involved in the biosynthesis of phase-variable oligosaccharide structures. In this paper we describe two additional chromosomal loci containing the sequence (CAAT)n. A survey of other *H. influenzae* isolates showed that there are from two to five regions containing the (CAAT)n sequence per isolate.

**MATERIALS AND METHODS**

**Bacterial strains.** The *H. influenzae* isolates used in hybridization studies included serotype b strains from each of the two phylogenetic lineages based on patterns of multilocus enzyme electrophoresis (17), strain Eagan (= RM.153) (lineage 1) and strain RM.926 (lineage 2). Other *H. influenzae* strains from our collection were strains RM.1152 (serotype a), RM.7422 (serotype c), RM.718 (serotype d), RM.6155 (serotype c), and RM.1137 (serotype f). Strain RM.6033 is a non-encapsulated *H. influenzae* strain obtained from a patient with 3304
salpingitis. Strain F.3031 (H.influenzae subtype aegyptius) was isolated from a patient with Brazilian purpuric fever; and strain F.3054 (H. influenzae subtype aegyptius) was isolated from a patient with conjunctivitis; both of these strains were provided by the Brazilian Purpuric Fever Study Group, Centers for Disease Control, Atlanta, Ga. Strain RM.3028 is an isolate of Haemophilus parainfluenzae. Strain RM.7004 (= 760705) (serotype b) has been described previously by van Alphen et al. (27). Strain RM.3226 is a Nesi seria meningitidis cerebrospinal fluid isolate. N. gonorrhoea P.9-17 was provided by J. E. Heckels, Southampton University, Southampton, United Kingdom.

Bacteriophage lambda clones were propagated in Escherichia coli Q358 (hsdR supE tonA) (11). M13 bacteriophages were grown in E. coli TG1 [Δlac-pro supE thi hsdR5(F' traD36 proA*B' lacF' lacZAM15)] T. J. Gibson, Ph.D. thesis, University of Cambridge, Cambridge, England, 1984). Plasmid constructions were maintained in E. coli DH5α (ΔlacZYA-argF)U169 φ80delZAM15 endA1 hsdR17 supE44 thi-l recA1 gyrA96 relA1) (10).

Culture conditions. H. influenzae strains were grown in brain heart infusion broth (Oxoid, Ltd., London, United Kingdom) supplemented with 2 μg of NAD (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) per ml and 10 μg of hemin (Sigma Chemical Co., St. Louis, Mo.) per ml; this supplemented broth is referred to below as sBHI broth. Brain heart infusion agar plates were prepared with 1% agar and were supplemented with 10% Levinthal base (sBHI agar) (1). Neisseria strains were grown on clear typing agar (26) and were physically removed from plates for DNA preparation. E. coli strains were grown on Luria agar to which kanamycin (40 μg ml⁻¹) or ampicillin (50 μg ml⁻¹) or both were added.

Colony immunoblotting. MAbs directed against H. influenzae LPS oligosaccharide epitopes have been described previously (12, 28). Additional anti-LPS murine MAbs, MAAb A1-12E-5G-8Es and A3-2F-10A-6b, were provided by A. Lindberg. For colony immunoblotting, a single colony of H. influenzae was grown to mid-log phase in sBHI broth containing (if necessary) kanamycin (10 μg ml⁻¹), and 500 to 1,000 of these organisms were grown on sBHI agar containing kanamycin (100 μg ml⁻¹). The technique used for colony immunoblotting has been described previously (28).

Construction of plasmids and H. influenzae mutants. Construction of the genomic library of strain RM.7004 DNA cloned into λEMBL3 has been described previously (28). DNA for subcloning was obtained by fractionating endonuclease-digested DNA from recombinant EMBL3 phage on 1% agarose gels in Tris-borate buffer. DNA bands were excised from the gels and purified by using Gene Clean (Stratagene) or an electroelution device (International Biotechnology, Inc., New Haven, Conn.) before ligation into the cloning sites of the vectors (pUC18, pUC19, pBluescript, M13mp18, and M13mp19) (18). Recombinants were distinguished from nonrecombinants by their failure to hydrolyze 5-bromo-4-chloro-3-indolyl β-D-galactoside (Bethyl Research Laboratories, Inc., Gaithersburg, Md.). plasmid DNA was prepared by alkaline lysis (5). Single-stranded DNA templates were prepared from recombinant M13 phage by using the polyethylene glycol precipitation method (22).

For deletion-insertion mutagenesis, plasmids containing lic2 or lic3 sequences were digested within the insertion in order either to linearize the plasmid (insertion mutant) or to excise a fragment within the insertion (deletion mutant). Sticky ends were made blunt by using the Klenow fragment of E. coli polymerase 1 (Amersham Corp., Arlington Heights, Ill.). The plasmid was ligated in the presence of a blunt-ended kanamycin resistant cartridge derived from Tn903 (19). The plasmid construction was linearized with SalI and added to competent strain RM.7004 cells (6). Colonies resistant to kanamycin (10 μg ml⁻¹) were selected on sBHI agar and then screened for sensitivity to ampicillin (20 μg ml⁻¹). Total cellular DNA was extracted as described previously (8).

DNA sequencing and sequence analysis. The nucleotide sequence was determined by the dyeoxy chain-termination method (23), using a DNA sequencing kit containing a modified T7 polymerase (Sequenase; United States Biochemical Corp.). DNA was labeled by including [α-32P]dATP (Amersham) in the reaction mixture (2). Plasmids were sequenced by using the method of Hettari and Sakaki (8). Data were analyzed by using the software written by Staden (25).

DNA hybridization studies. Endonuclease fragments of DNA were transferred from agarose gels to nylon filters for Southern hybridization (24) with probes labeled by nick translation in the presence of [α-32P]dCTP (21). The specific probes included a 3.2-kilobase (kb) BamHI-HindIII fragment of lic1, a 3.1-kb XbaI fragment of lic2, and a 0.4-kb BglII-Sau3AI fragment of lic3. High-stringency hybridization was carried out in 45% formamide at 42°C, which allowed for 11% base pair mismatches. Low-stringency hybridization was carried out in 20% formamide at 37°C, which allowed for 34% base pair mismatches (15).

The oligonucleotide 5'-CAATCAATCAATCAATCAAT-3' was prepared by using a model 381 DNA synthesizer (Applied Biosystems). This oligonucleotide was 5'-end labeled with [γ-32P]ATP by using T4 polynucleotide kinase. DNA fixed on nylon filters was hybridized with this probe at 37°C for 1 h as described previously (16). Filters were washed in 6 × SSC (1 × SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 45°C (5°C below the calculated melting temperature).

Genomic mapping. High-molecular-weight strain Eagan DNA was prepared in low-melting-point agarose (Bethesda Research Laboratories, Inc., Gaithersburg, Md.), digested with either RsrII, NaeI, EaeI, or Smal, and separated by contour-clamp homogenous electric field gel electrophoresis (5) in 1% agarose (Sigma). Concatamers of bacteriophage λ DNA were used as size markers. The DNA was transferred to nylon filters and probed with nick-translated plasmids containing lic sequences. A genomic map was constructed by using the method of Butler and Moxon (unpublished data).

RESULTS

Identification and characterization of other (CAAT)n-containing loci. When whole-cell, genomic, endonuclease-digested DNA from strain RM.7004 was probed with the cloned DNA fragments from lic1 (Fig. 1), hybridization to two other chromosomal fragments was observed. This result was obtained only when the DNA sequence of the probe included the multiple repeats of CAAT found at the 5' end of lic1. To test whether the homology was due to the CAAT repeats, endonuclease-digested total genomic DNA from H. influenzae RM.7004 was probed with a 32P-labeled oligonucleotide having the sequence 5'-CAATCAATCAATCAAT-3'. We identified two BglII fragments containing the sequence (CAAT)n, in addition to lic1 (5.5 kb) (Fig. 2, lane
Further evidence for the presence of three genomic regions with (CAAT)n sequences was obtained by performing additional hybridizations in which strain RM.7004 DNA was digested with different endonucleases. The two additional regions were cloned from a λEMBL3 genomic library of RM.7004 DNA which was screened by using radiolabeled oligonucleotide (CAAT)n. In order to identify lambda clones containing the two additional regions, 11 plaques which hybridized to the probe were selected, and each was amplified to allow preparation of sufficient DNA for preliminary mapping with restriction endonucleases. On the basis of endonuclease digestion patterns, the λEMBL3 clones included sequences which were from three distinct regions of the chromosome, one of which corresponded to lic1. The two novel loci were designated lic2 and lic3. Restriction maps of each of these regions, each spanning about 5 kb of the genome, are shown in Fig. 1.

The positions of the (CAAT)n sequences in lic2 and lic3 were identified by Southern hybridization of endonuclease digests of the λEMBL3 clones, using the (CAAT)n oligonucleotide probe. The (CAAT)n sequence in lic2 was contained on a 0.4-kb BglII-XbaI fragment, and the (CAAT)n sequence in lic3 was found on a 0.9-kb HindIII-BglII fragment (Fig. 1).

These regions were subcloned into plasmid vectors. Smaller restriction fragments containing (CAAT)n were subcloned into either pUC or M13 vectors to obtain their nucleotide sequences. Analysis of these sequences (Fig. 3) revealed 16 and 22 tandem repeats of CAAT in lic2 and lic3, respect-

FIG. 1. Restriction endonuclease maps of chromosomal loci containing lic1, lic2, and lic3. The locations of the (CAAT)n sequence are indicated by hatched boxes. Each locus begins 5' with respect to the (CAAT)n sequence and extends 3' as indicated by the solid bars (based on nucleotide sequencing). Mutagenesis experiments indicated the approximate locations of sequences required for expression of epitopes specified by MAb e.g., MAb 6A2. Endonuclease sites are labeled as follows: B, BamHI; G, BglII; C, Clal; E, EcoRI; R, EcoRV; H, HindIII; S, Sau3AI; X, XbaI. A1, MAb A1-12E-5G, 8E; A3, MAb A3-2F-10A-6b; kbp, kilobase pair.

FIG. 2. (A) Homology to oligonucleotide (CAAT)n in electrophoretically separated DNAs from isolates digested with BglII (lanes 1 through 12) or EcoRI (lanes 13 and 14). (B) Schematic representation showing genomic fragments (bars) which correspond to bands in panel A. Bars are labeled 1, 2, or 3 to indicate homology to either lic1, lic2, or lic3. Unlabeled bars lack homology to lic1-, lic2-, and lic3-specific probes. The preparation in lane 12 showed no homology to the (CAAT)n probe but hybridized with the lic2 probe at the position indicated. Lane 1, strain RM.7004 (serotype b); lane 2, strain RM.1152 (serotype a); lane 3, strain Eagan (serotype b, lineage 1); lane 4, strain RM.926 (serotype b, lineage 2); lane 5, strain RM.7422 (serotype c); lane 6, strain RM.1118 (serotype d); lane 7, strain RM.6135 (serotype e); lane 8, strain RM.1137 (serotype f); lane 9, strain RM.6033 (nontypable); lane 10, strain F.3031 (subtype aegyptius, Brazilian purpuric fever); lane 11, strain F.3054 (subtype aegyptius, non-Brazilian purpuric fever); lane 12, strain RM.3028 (H. parainfluenzae); lane 13, strain RM.3226 (N. meningitidis); lane 14, strain P.9-17 (N. gonorrhoeae). The sizes of the fragments (in kilobases) are indicated on the left.
sequence in E. coli (20). Following the last tetrapeptide (SINQ), there is no similarity in the deduced amino acid sequences for *lic1*, *lic2*, and *lic3* for those portions of the proteins whose sequences were established (Fig. 3).

**Function of *lic2* in LPS expression.** The role of *lic2* in oligosaccharide expression was defined by mutagenesis. Plasmid DNA in which an internal fragment of an insertion derived from *lic2* was replaced by a kanamycin resistance gene was introduced into the chromosome of strain RM.7004 by transformation. Southern hybridization confirmed that the resulting transformant contained the resistance gene in the chromosome and had the predicted deletion (data not shown). The LPS phenotype of the deletion-insertion mutant was assessed by colony immunoblotting. *lic2* was contained with a 3.5-kb *BglII*-Xbal region (Fig. 1), 3.2 kb of which was located downstream from the (CAAT)n sequence. The deletion of sequences either 5′ or 3′ with respect to this 3.5-kb region had no effect on the expression of any of the epitopes recognized by the six currently available MAb. The 3.5-kb *lic2* region contained two internal Xbal fragments. Deletion of both of these fragments (3.1 kb) resulted in constitutive loss of reactivity with MAb **4C**, **5G**, A1-12E-5G-8E and A3-2F-10A-6B (mutant RM.7004-X) (Table 1). Deletion of the downstream 1.3-kb Xbal fragment in mutant RM.7004-H eliminated reactivity with MAb A1-12E-5G-8E but had no effect on the expression of the other epitopes. On the basis of these data and the previously reported data for mutagenesis of *lic1* (mutant RM.7004-RV2) (29) (Table 1), it was possible to identify *lic1* required for the expression of the six phase-variable LPS epitopes (Fig. 1).

The organization of *lic1* and *lic2* is similar. Both loci have a (CAAT)n sequence upstream from the genes required for expression of several oligosaccharide structures, although *lic1* and *lic2* determine expression of different variable LPS epitopes.

The kanamycin resistance gene was inserted into a unique *BglII* site downstream from, but within the same open reading frame as, the CAAT repeats in *lic3* (Fig. 1). Kanamycin-resistant mutant RM.7004-S appeared to be unchanged from the parent strain with respect to its reactivity to each of the six MAb that recognized phase-variable epitopes (Table 1). Thus, all known phase-variable LPS structures could be attributed to either *lic1* or *lic2*. So far, it has not been possible to demonstrate a role for *lic3* in LPS expression.

**Location of *lic1*, *lic2*, and *lic3* on a genomic map.** An endonuclease map of the genome of strain Eagan was constructed. This map was obtained by using four endonucleases (RsuRI, EagI, NaeI, and SmaI) which cleaved the A-T-rich *H. influenzae* chromosome infrequently (Butler and

---

**TABLE 1. Reactivity of *H. influenzae* mutants with MAbs to phase-variable LPS epitopes**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Mutated locus</th>
<th>Reactivity with MAb*</th>
</tr>
</thead>
<tbody>
<tr>
<td>RM.7004</td>
<td>12D9</td>
<td>PV PV PV PV PV PV</td>
</tr>
<tr>
<td>RM.7004-RV2</td>
<td>lic1</td>
<td>– – PV PV PV PV</td>
</tr>
<tr>
<td>RM.7004-X</td>
<td>lic2</td>
<td>PV PV – – – –</td>
</tr>
<tr>
<td>RM.7004-H</td>
<td>lic2</td>
<td>PV PV PV PV PV PV</td>
</tr>
<tr>
<td>RM.7004-S</td>
<td>lic3</td>
<td>PV PV PV PV PV PV</td>
</tr>
</tbody>
</table>

* Reactivity with anti-LPS-directed MAbs was determined by colony immunoblotting. Mutants either retained expression and phase variation of an epitope (PV) or were nonreactive (–). Mutant RM.7004-RV2 has been described previously (31).

---

**FIG. 3.** Nucleotide and amino acid sequences of the *lic* loci in the (CAAT)n region. For each *lic* locus (*lic1*, *lic2*, and *lic3*) nucleotide sequences around the repeat region [(CAAT)n] are shown. For each locus, the sequence given is for n + 1 repeats, the value of n having been determined by nucleotide sequencing. Repeat regions are underlined. Deduced, putative translation products (amino acid sequences) are shown for n + 1 and n – 1 repeats. For the latter, which differ by having two fewer repeats (n – 1), the deleted repeats are represented by dashes. These are the only deduced products which give full-length polypeptides because of the presence of stop codons in the other reading frames after the repeat regions.

Open reading frames, which included the CAAT repeats, were found in all three loci. The (CAAT)n region encoded multiple copies of the tetrapeptide serine-isoleucine-glutamine-asparagine (SINQ). As has been shown to occur in *lic1* (29), variations in the number of repeats placed nearby initiation codons in or out of phase for translation. In *lic2*, there is a single ATG separated by 11 base pairs from the first CAAT copy; in *lic3*, two ATGs are located in different translational phases, separated by 1 and 15 base pairs, respectively, from the initial repeat. The putative translational products of *lic1* and *lic3* (n – 1 repeats) should be similar in the sequence encoded upstream from the first CAAT unit. None of the loci had identifiable Shine-Dalgarno sequences upstream from their initiation codons. However, the sequence (CAAT)n provides a significant 5-base-pair match with a downstream 16S rRNA-binding consensus

---

Vol. 172, 1990
Moxon, unpublished data). Strain Eagan is a serotype b strain that is closely related to strain RM.7004 based on the results of multilocus enzyme electrophoresis (17). High-molecular-weight genomic DNA embedded in agarose was digested by one of the endonucleases and separated by contour-clamp homogenous electric field gel electrophoresis (5). Southern hybridization in which lic1-, lic2-, and lic3-specific probes were used revealed that all three loci were situated on a 215-kb EagI fragment. More detailed mapping of this fragment showed that lic1 was on a 50-kb EagI-NaeI fragment, lic2 on the adjacent 30-kb NaeI-SmaI fragment, and lic3 was on the adjacent 105-kb SmaI-NaeI fragment. From a map of the regions flanking lic1, lic2, and lic3 the minimum distance that separated the loci was shown to be 14 kb. Therefore, all three loci are located within a region which constitutes 8.3% of the genome, although they are separated by intervening stretches of DNA of unknown function.

(CAAT)n sequences are present in other H. influenzae strains. Representative Haemophilus strains were examined for the presence of (CAAT)n sequences. Whole genomic DNA digested with BgIII was probed with 32P-labeled oligonucleotide (CAAT)n. The isolates contained from two to five regions with this sequence (Fig. 2). There was no hybridization to the DNA of a strain of H. parainfluenzae.

Duplicate filters were probed with sequences specific for lic1, lic2, or lic3 under conditions of high stringency. Both lic1 and lic2 were found in each of the H. influenzae strains. The biotype aegyptius isolate which was obtained from a patient with Brazilian purpuric fever had two copies of the lic1 sequence (confirmed by digestion with different endonucleases). The lic2-specific probe hybridized to the H. parainfluenzae genome, although there was no (CAAT)n sequence in this strain. In contrast, a lic3-specific probe hybridized to some, but not all, H. influenzae strains. There were several regions containing (CAAT)n which did not hybridize to the lic1-, lic2-, and lic3-specific probes. Thus, although the (CAAT)n sequences were common to the H. influenzae strains, the patterns of hybridization were variable, as evidenced by the restriction fragment length polymorphisms.

N. gonorrhoeae and N. meningitidis have shown to share several LPS epitopes with H. influenzae, including structures which undergo phase variation (Virji et al., unpublished data). However, DNAs from gonococcal and meningococcal isolates did not hybridize when they were probed with either the (CAAT)n oligonucleotide or the lic1-, lic2-, and lic3-specific probes, even under conditions of low stringency.

DISCUSSION

In previous studies we described the chromosomal locus lic1, which controls both expression and phase variation of at least two core LPS structures of an H. influenzae serotype b strain, strain RM.7004 (28, 29). In this study, we found that RM.7004 has two additional chromosomal loci (lic2 and lic3) which also contain multiple tandem repetitions of the tetranucleotide CAAT. Furthermore, these tetramer repeats are situated 5' with respect to long open reading frames so that the overall organization is similar to that of lic1. lic2 controls the expression of LPS core oligosaccharide structures which exhibit phase variation and which are different from those assigned to the control of lic1. Thus, in RM.7004, lic1 controls the expression of the MAb 6A2-specific epitope, which in turn depends on the presence of the MAb 12D9-specific epitope (lic1). Similarly, lic2 controls the MAb 5G8-specific epitope, which depends on the presence of the epitope specified by MAb 4C4 (lic2). However, these functionally separate control mechanisms (in lic1 and lic2) confer on strain RM.7004 the ability to switch on or off different LPS antigens independently. We also found that from two to five loci containing (CAAT)n exist in other H. influenzae strains.

The organization of lic2 and lic3 is similar to the organization of lic1, although the function of lic3 is unknown. In each locus repetitive DNA sequences are found at the 5' end, which could mediate phase variation through a translational mechanism similar to that described previously for lic1. Specifically, depending on the number of CAAT units, upstream ATGs would be either in or out of frame in each instance (a number of repeats corresponding to an out-of-frame situation is a consistent finding when these sequences are cloned and amplified in an E. coli host). In lic2 only one translational phase with an upstream ATG was found. This is consistent with the observation that epitopes controlled by lic2 are either on or off. In contrast, lic1 and lic3 have closely situated ATGs in two of three phases of translation, and in the case of lic1 we found that the (CAAT)n sequence is capable of determining both phase variation and differences in the level of expression of the relevant LPS epitope.

A detailed map of strain Eagan shows that in this strain the three (CAAT)n regions, lic1, lic2, and lic3, are located within a relatively small region representing less than 10% of the entire genome of approximately 2 × 106 kb. The close proximity of the (CAAT)n regions may facilitate coordinate regulation of phase variation. Indeed, it has been observed previously that a loss of lic1-controlled epitopes may be accompanied by a simultaneous loss of lic2 epitopes (28). Since the on-off translational switching of lic1 and presumably also lic2 is determined by different repetitive sequences of CAAT, coincident switching off of both lic1 and lic2 epitope expression would not be expected to occur frequently on a chance basis.

ACKNOWLEDGMENTS

This project was supported by the Meningitis Trust (United Kingdom), by Programme Grant 8525352 from the Medical Research Council, and by grant 85-04311P from the Swedish Board for Technical Development (to A.A.L.).

We thank Eve Hansen for supplying H. influenzae LPS Mabs. J.N.W. is a Fellow of the Institute for Pediatric Service, Johnson & Johnson Baby Products Co. We thank Sheila Hayes for her invaluable expert help in the preparation of the manuscript.

LITERATURE CITED


