Journal of Bacteriology, June 1996, p. 3339–3341
0021-9193/96/$04.00+0
Copyright © 1996, American Society for Microbiology
Vol. 178, No. 11

Molecular Cloning of the *Haemophilus influenzae gmhA* (*lpcA*) Gene Encoding a Phosphoheptose Isomerase Required for Lipooligosaccharide Biosynthesis

JOANNA S. BROOKE and M. A. VALVANO*

Department of Microbiology and Immunology, University of Western Ontario, London, Ontario N6A 5C1, Canada

Received 26 January 1996/Accepted 20 March 1996

We have determined that gene HI#1181 of *Haemophilus influenzae* is a homolog of *Escherichia coli gmhA* (previously designated *lpcA*) (J. S. Brooke and M. A. Valvano, J. Biol. Chem. 271:3608–3614, 1996), which encodes a phosphoheptose isomerase catalyzing the first step of the biosynthesis of ADP-L-glycero-D-manno-heptose. Mutations in this gene are associated with a heptoseless core lipopolysaccharide which determines an increased outer membrane permeability to hydrophobic compounds. The cloned *H. influenzae gmhA* restored the synthesis of a complete core in the *gmhA*-deleted *E. coli* strain χ711. Amino acid sequence comparisons of the GmhA proteins of *E. coli* and *H. influenzae* with other proteins in the databases revealed the existence of a novel family of phosphosugar aldo-keto isomerases.

Haemophilus influenzae is a common inhabitant of the nasopharynx and can be a causative agent of otitis media and complications of chronic bronchitis (27). Capsulated *H. influenzae* type b strains are responsible for invasive diseases in infants, such as meningitis, epiglottitis, and pneumonia (27). Production of capsular polysaccharide is a major virulence factor in this bacterium (16), but also other factors have been associated with virulence, including the lipooligosaccharide (LOS) (4, 12, 15, 29, 30).

H. influenzae LOS is composed of lipid A and a heterogeneous oligosaccharide, and it is analogous to the lipopolysaccharide (LPS) of enteric bacteria, except that the LOS lacks O-specific polysaccharide. The analysis of the structure of the LOS revealed a unique inner core structure with variable branched chains (18). The inner structure contains 3-deoxy-D-*manno*-octulosonic acid (KDO) and at least three glycero-mannoheptose residues (2), and it is practically identical to the inner core of enteric bacteria.

Presumably, genes with conserved functions are involved in directing the synthesis and assembly of the inner core in most, if not all, gram-negative bacteria. One of these genes, recently found to be conserved in enteric bacteria, is *lpcA* (26), which encodes a phosphoheptose isomerase catalyzing the conversion of sedoheptulose-7-phosphate into D-glycero-D-manno-heptose-7-phosphate (3). This is the first step committed to the biosynthesis of the ADP-L-glycero-D-mannoheptose precursor for the formation of the inner core (7).

lpcA stands for lipopolysaccharide core biosynthesis, and although appropriate, this designation is too general, especially when the numerous designations used for other lipid A core synthesis genes are considered. There is a concerted effort by our and other laboratories working on genetics and biosynthesis of lipopolysaccharides and other surface carbohydrates to systematize the nomenclature of biosynthetic genes (22). All these authors have agreed to adopt *gmh* for the designation of genes involved in the synthesis of ADP-glyceromannoheptose (22); consequently, *gmhA* instead of *lpcA* will be used throughout this report.

* Corresponding author. Phone: (519) 661-3996. Fax: (519) 661-3499. Electronic mail address: mvalvano@uwo.ca.

The study of the genes required for LOS biosynthesis has been difficult because of the lack of genetic tools for *H. influenzae* (14). Since D-glycero-D-mannoheptose has been observed in the core regions of LOS of *Haemophilus ducreyi*, LOS of *H. influenzae* (24), and LPS from enteric bacteria, it would be a reasonable approach to isolate *Haemophilus* genes responsible for the biosynthesis of this region by complementing *Salmonella* or *Escherichia coli* LPS mutants. This approach has proven to be successful, as evidenced by the recent cloning of the *H. influenzae rfaE* gene, believed to be involved in the synthesis of ADP-glyceromannoheptose, by complementation of a *Salmonella rfaE* mutant (14). Furthermore, the recent elucidation of the complete DNA sequence of the *H. influenzae* genome (8) has greatly facilitated direct comparisons of gene and protein sequences with those of other species, making it possible to identify the functions of otherwise unknown *H. influenzae* genes.

In this paper, we report the identification, molecular cloning, and expression of the *gmhA* homolog of *H. influenzae* which was previously considered a hypothetical gene with unknown function (8).

We have recently cloned the *gmhA* gene from *E. coli* and found it to encode a polypeptide of 192 amino acids (3). A search of the *H. influenzae* genomic database (http://www.tigr.com) of translated proteins revealed a polypeptide of 194 amino acids corresponding to the product of gene HI#1181, which exhibited 74% amino acid sequence identity with the *E. coli* GmhA (Fig. 1). Because of this strong conservation, we sought to clone this gene and used it to complement a *gmhA* mutant of *E. coli.*

Synthetic oligonucleotide primers 5′CTTGATTGAAATCAACCGCAC and 5′CTTGAAAAATCCCCAACT flanking the 5′ and 3′ ends of HI#1181, respectively, were used in a PCR mixture containing chromosomal DNA of *H. influenzae* MAP7 (21). The PCR was carried out with a Hybaid Omnigene Temperature Cycler (Interscience, Markham, Ontario, Canada) with *Pwo*I DNA polymerase (Boehringer Mannheim, Dorval, Quebec, Canada). The PCR consisted of one cycle of denaturation at 95°C for 1 min; 28 cycles of denaturation at 95°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 2 min; and one cycle of denaturation at 95°C for 1 min, an-

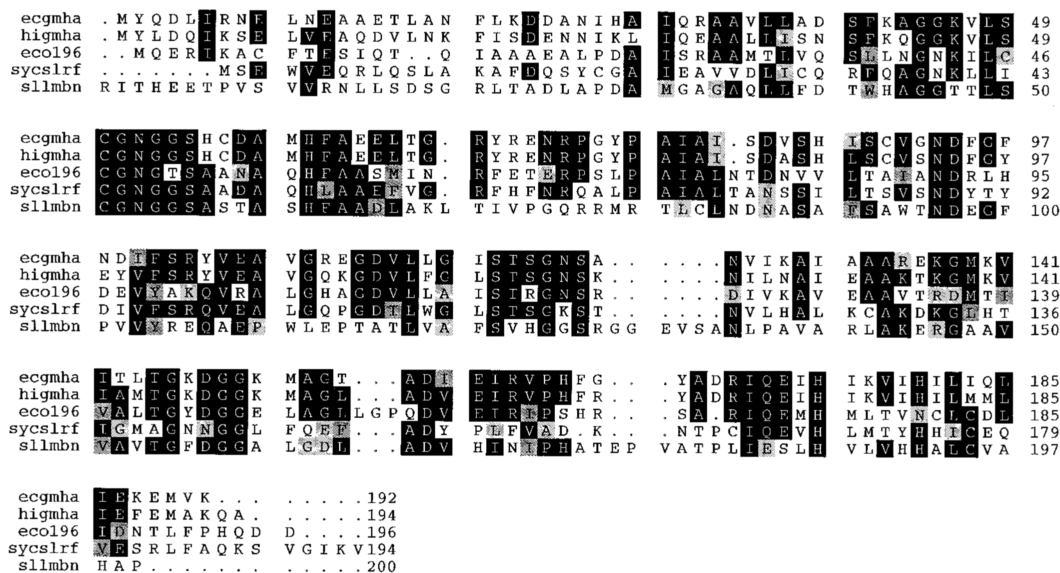

FIG. 1. Amino acid sequence comparisons of the GmhA proteins of *E. coli* and *H. influenzae* with other proteins in databases. Sequences identified by a BLAST (1) search were aligned with the PILEUP program from the University of Wisconsin Genetics Computer Group package, version 8.1 (6). Protein sequences are abbreviated as follows: ecgmha, *E. coli* GmhA (accession no. U32590); higmha, *H. influenzae* GmhA (accession no. U32797); eco196, *E. coli* o196, a hypothetical protein in the 67- to 76-min region of the chromosome (accession no. U18997); sycslrf, a *Synechocystis* sp. hypothetical protein (11) (accession no. D64004); sllmbn, *Streptomyces lincolnensis* LmbN (accession no. X79146). Shaded areas indicate identical amino acids; crosshatched areas indicate conserved amino acids. The comparisons involve the entire lengths of all the proteins except for LmbN (see the text).

nealing at 55°C for 1 min, and extension at 72°C for 10 min. The amplified fragment of ca. 760 bp was treated with T4 polynucleotide kinase (31 U) (Pharmacia Canada Inc., Baie d'Urfe, Quebec, Canada) for 30 min at 37°C, and purified from a 0.7% agarose gel with a QIAquick gel extraction kit (Qiagen, Chatsworth, Calif.). This fragment was ligated into the *Sma*I site of the vector pMAV3 (28) and transformed by electroporation into *E. coli* χ711 (F$^-$ *leu-4* φ^r *proAB118* Str^r T3^r *arg-35* T6^r). This strain was isolated by Curtiss et al. (5) and has a chromosomal deletion that removes the *gmhA* gene as well as the proline synthesis genes *proAB* (3, 5). Since deletion of *gmhA* results in a heptoseless LPS, *E. coli* χ711 cells are unable to grow on Luria broth supplemented with 50 μg of novobiocin per ml (3). *E. coli* χ711 transformants which were able to grow on plates containing novobiocin indicated that the cloned PCR fragment complemented the mutant phenotype. Plasmids isolated from these colonies contained a ca. 760-bp DNA insert, and one of them, designated pJB24, was used for DNA sequencing to confirm that the HI#1181 gene was indeed cloned. An open reading frame identical to HI#1181 was identified by DNA sequencing carried out with double-stranded DNA and the same primers used for the PCR (data not shown).

To confirm that the novobiocin-resistant phenotype of χ711 (pJB24) was due to the restoration of the *gmhA* function, we examined the LPS profile of this strain by Tricine sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Figure 2 shows that the length of the LPS extracted from χ711 (lane B) is shorter than that produced by the isogenic parental strain χ705 (lane A) and identical to that produced by strain D31m4 (lane D). Strain D31m4 (*rfa-229 rfa-230*) produces a heptoseless LPS which contains only lipid A and KDO (20). A comparison of the core LPS profiles of strains χ705 and χ711 (pJB24) showed that this plasmid restored the core LPS defect of *E. coli* χ711 (Fig. 2, lanes A and C). Therefore, we conclude that pJB24 carries the *gmhA* homolog of *H. influenzae.*

While this report was in preparation, Preston et al. reported the identification of the *isn* gene encoding an unknown function involved in the synthesis of LOS (19). Previous work has also shown that *isn* mutants formed an incomplete LOS containing only phosphorylated KDO and lipid A (9) and were associated with serum sensitivity and loss of virulence in a rat model (30). Since the *isn* and *gmhA* DNA sequences are identical, we conclude that they correspond to the same gene, and on the basis of our biochemical characterization of the *E. coli gmhA* (3) as well as on the complementation studies reported here, we demonstrate that the *H. influenzae* gene encodes a phosphoheptose isomerase. We propose that the designation *isn* (insertion sixty-nine) be replaced by *gmhA*, which gives the appropriate functional connotation to this gene.

The percent G+C content of *H. influenzae gmhA* was 36.2%, similar to the reported values for percent G+C content of the *H. influenzae* genome (25), and the codon usage was typical for *H. influenzae* genes. This suggested that the *gmhA* gene has not recently evolved from horizontal transfer. Like the *gmhA* gene of *E. coli* K-12 (3), *H. influenzae gmhA* is physically apart from other genes used in LOS biosynthesis, such as the *lsg* and *lic* loci (8). Also, like the *E. coli* counterpart (3), *H. influenzae gmhA* is not transcriptionally linked to adjacent genes (19).

A comparison of the amino acid sequences of the GmhA

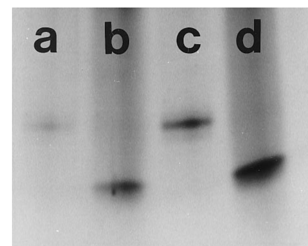

FIG. 2. Tricine SDS-PAGE analysis of lipid A core oligosaccharides. Lanes: A, *E. coli* χ705; B, *E. coli* χ711; C, *E. coli* χ711(pJB24); D, *E. coli* D31m4 (*rfa-229 rfa-230*).

polypeptides from both *E. coli* and *H. influenzae* with other proteins in databases (Fig. 1) revealed strong similarities with uncharacterized open reading frames from *E. coli* and *Synechocystis* sp. and with the LmbN protein of *Streptomyces lincolnensis* (64, 61, and 53% amino acid sequence identity, respectively). These proteins also displayed very similar hydropathy profiles (data not shown; 13). Except for LmbN, which is a polypeptide of 275 amino acids, the other four polypeptides ranged from 192 to 196 amino acids in length. The amino acid conservation extended throughout the entire protein sequence, but it was especially strong in the central region of these proteins. We have previously shown that at least part of this region in the *E. coli* GmhA protein has similarities with a group of glutamine:fructose-6-phosphate amidotransferases (3), which are also aldo-keto isomerases (23). The functions of the *E. coli* polypeptide o196 and the *Synechocystis* polypeptide are not known, but LmbN is involved in the biosynthesis of the antibiotic lincomycin (17), which requires an isomerization step of an eight-carbon phosphosugar precursor (10); therefore, it is likely that LmbN is also an isomerase. We thus propose that all these proteins define a novel family of phosphosugar isomerases.

We thank R. J. Redfield and C. Whitfield for the gift of *H. influenzae* chromosomal DNA.

This work was supported by a grant of the Medical Research Council of Canada to M.A.V. J.S.B. was supported by an Ontario Graduate Scholarship.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–441.
2. **Borrelli, S., O. Hegedus, D. H. Shaw, P.-E. Jansson, and A. A. Lindberg.** 1995. The tetrasaccharide L-α-D-heptose1→2-L-α-D-heptose1→3-L-α-D-heptose1→(3-deoxy-D-*manno*-octulosonic acid) and phosphate in lipid A define the conserved epitope in *Haemophilus* lipopolysaccharides recognized by a monoclonal antibody. Infect. Immun. **63:**3683–3692.
3. **Brooke, J. S., and M. A. Valvano.** 1996. Biosynthesis of inner core lipopolysaccharide in enteric bacteria identification and characterization of a conserved phosphoheptose isomerase. J. Biol. Chem. **271:**3608–3614.
4. **Cope, L. D., R. Yogev, J. Mertsola, J. C. Argyle, G. H. McCracken, Jr., and E. J. Hansen.** 1990. Effect of mutations in lipooligosaccharide biosynthesis genes on virulence of *Haemophilus influenzae* type b. Infect. Immun. **58:** 2343–2351.
5. **Curtiss, R., J. Charamella, D. R. Stallions, and J. A. Mays.** 1968. Parental functions during conjugation in *Escherichia coli* K-12. Bacteriol. Rev. **32:** 320–348.
6. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acid Res. **12:**387–395.
7. **Eidels, L., and M. J. Osborn.** 1974. Phosphoheptose isomerase, first enzyme in the biosynthesis of aldoheptose in *Salmonella typhimurium*. J. Biol. Chem. **249:**5642–5648.
8. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L. I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Furhmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
9. **Helander, I. M., B. Lindner, H. Brade, K. Altmann, A. A. Lindberg, E. T. Rietschel, and U. Zähringer.** 1988. Chemical structure of the lipopolysaccharide of *Haemophilus influenzae* strain I-69 Rd−/B+: description of a novel deep-rough chemotype. Eur. J. Biochem. **177:**483–492.
10. **Howarth, G. B., D. G. Lance, W. A. Szarek, and J. K. N. Jones.** 1969. Syntheses related to the carbohydrate moiety in lincomycin. Can. J. Chem. **47:**75–79.
11. **Kaneko, T., A. Tanaka, S. Sato, H. Kotani, T. Sazuka, N. Miyajima, M. Sugiura, and S. Tabata.** 1995. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1Mb region from map positions 64% to 92% of the genome. DNA Res. **2:**153–166.
12. **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.
13. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
14. **Lee, N.-G., M. G. Sunshine, and M. A. Apicella.** 1995. Molecular cloning and characterization of the nontypeable *Haemophilus influenzae* 2019 *rfaE* gene required for lipopolysaccharide biosynthesis. Infect. Immun. **63:**818–824.
15. **Moxon, E. R., and D. J. Maskell.** 1992. *Haemophilus influenzae* lipopolysaccharide: the biochemistry and biology of a virulence factor, p. 75–96. *In* C. E. Hormaeche, C. W. Penn, and C. J. Smyths (ed.), Molecular biology of bacterial infection: current status and future perspectives. Cambridge University Press, Cambridge.
16. **Moxon, E. R., and K. A. Vaughn.** 1981. The type b capsular polysaccharide as a virulence determinant of *Haemophilus influenzae*: studies using clinical isolates and laboratory transformants. J. Infect. Dis. **143:**517–524.
17. **Peschke, U., H. Schmidt, H.-Z. Zhang, and W. Piepersberg.** 1995. Molecular characterization of the lincomycin-production gene cluster of *Streptomyces lincolnensis* 78-11. Mol. Microbiol. **16:**1137–1156.
18. **Phillips, N. J., M. A. Apicella, J. McLeod Griffis, and B. W. Gibson.** 1993. Structural studies of the lipooligosaccharides from *Haemophilus influenzae* type b strain A2. Biochemistry **32:**2003–2012.
19. **Preston, A., D. Maskell, A. Johnson, and E. R. Moxon.** 1996. Altered lipopolysaccharide characteristic of the I69 phenotype in *Haemophilus influenzae* results from mutations in a novel gene, *isn*. J. Bacteriol. **178:**396–402.
20. **Qureshi, N., K. Takayama, P. Mscagni, J. Honovich, R. Wong, and R. J Cotter.** 1988. Complete structural determination of lipopolysaccharide obtained from deep rough mutant of *Escherichia coli*. Purification by high performance liquid chromatography and direct analysis by plasma desorption mass spectrometry. J. Biol. Chem. **263:**11971–11976.
21. **Redfield, R. J.** 1991. *sxy-1*, a *Haemophilus influenzae* mutation causing greatly enhanced spontaneous competence. J. Bacteriol. **173:**5612–5618.
22. **Reeves, P. R., M. Hobbs, M. A. Valvano, K. Sanderson, M. Skurnik, C. Whitfield, D. Coplin, N. Kido, J. Klenna, C. Lee, D. Maskell, C. H. Raetz, and P. Rick.** Submitted for publication.
23. **Rose, I. A.** 1975. Mechanism of the aldose-ketose isomerase reactions. Adv. Enzymol. **43:**491–517.
24. **Schweda, E. K. H., A. C. Sundstrom, L. M. Eriksson, J. A. Jonasson, and A. A. Lindberg.** 1994. Structural studies of the cell envelope lipopolysaccharides from *Haemophilus ducreyi* strains ITM 2665 and ITM 4747. J. Biol. Chem. **269:**12040–12048.
25. **Smith, H. O., J. F. Tomb, B. A. Daugherty, R. D. Fleischmann, and J. C. Venter.** 1995. Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. Science **269:**538–540.
26. **Tamaki, S., T. Sato, and M. Matsuhashi.** 1971. Role of lipopolysaccharides in antibiotic resistance and bacteriophage adsorption of *Escherichia coli* K-12. J. Bacteriol. **105:**968–975.
27. **Turk, D. C.** 1984. The pathogenicity of *Haemophilus influenzae*. J. Med. Microbiol. **18:**1–16.
28. **Yao, Z., H. Liu, and M. A. Valvano.** 1992. Acetylation of O-specific lipopolysaccharides from *Shigella flexneri* 3a and 2a occurs in *Escherichia coli* K-12 carrying cloned *S. flexneri* 3a and 2a *rfb* genes. J. Bacteriol. **174:**7500–7508.
29. **Zwahlen, A., L. G. Rubin, C. J. Connelly, T. J. Inzana, and E. R. Moxon.** 1985. Alteration of the cell wall in *Haemophilus influenzae* type b by transformation with cloned DNA: association with attenuated virulence. J. Infect. Dis. **152:**485–492.
30. **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.