

Molecular Cloning, Nucleotide Sequence, and Characterization of a 40,000-Molecular-Weight Lipoprotein of *Haemophilus somnus*†

MICHAEL THEISEN,^{1,‡*} CLEMENT R. RIOUX,¹ AND ANDREW A. POTTER^{1,2}

*Canadian Bacterial Disease Network² and Veterinary Infectious Disease Organization,¹
University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 0W0*

Received 8 October 1991/Accepted 12 December 1991

A gene of *Haemophilus somnus* encoding the major 40,000-molecular-weight antigen (LppA) was cloned on a 2-kb *Sau3AI* fragment. The nucleotide sequence of the entire DNA insert was determined. One open reading frame, encoding a 247-residue polypeptide with a calculated molecular weight of 27,072, was identified. This reading frame was confirmed by sequencing the fusion joint of two independent *lppA::TnphoA* gene fusions. The 21 amino-terminal amino acids of the deduced polypeptide showed strong sequence homology to the signal peptide of secreted proteins, and the sequence Leu-Leu-Ala-Ala-Cys at the putative cleavage site is identical to the consensus cleavage sequence of lipoproteins from gram-negative bacteria. The presence of the lipid moiety on the protein was shown by incorporation of radioactive palmitic acid into the natural *H. somnus* protein. Palmitic acid could also be incorporated into the recombinant protein in *Escherichia coli*. Synthesis of the mature LppA lipoprotein was inhibited by globomycin, showing that cleavage of the signal peptide is mediated by signal peptidase II in both organisms. By using site-directed mutagenesis, the cysteine residue at the cleavage site was changed to glycine. Radiolabelled palmitate was not incorporated into the mutated protein, showing that lipid modification occurs at the Cys-22 residue.

Haemophilus somnus is a gram-negative bacterium which causes a number of disease syndromes in cattle. It is commonly associated with thromboembolic meningoencephalitis, septicemia, arthritis, and pneumonia (7, 16, 19). In North America, current vaccines are based either on killed whole cells or on a protein fraction enriched in outer membrane proteins. Only the latter offers significant protection against *H. somnus*-induced pneumonia in an experimental *H. somnus* challenge model (14). The outer membrane of *H. somnus* consists predominantly of proteins migrating at 43,000-molecular-weight (43K) and 40K bands on sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE). Of these, only a 40K protein reacts with convalescent-phase serum (8, 11).

An increasing number of bacterial antigens are being identified as lipoproteins (1, 5, 13, 18, 28, 34). These proteins are localized in the envelope of the cell and are therefore exposed to the host's immune system. It has been shown that the murein lipoprotein from the outer membrane of *Escherichia coli* acts as a potent activator of murine lymphocytes, inducing both proliferation and immunoglobulin secretion (4, 25). The lipoprotein part responsible for the biological activity was shown to reside in the N-terminal fatty acid-containing region (3). Recent results using synthetic lipopeptides show that even short peptides containing two to five amino acids covalently linked to palmitate possess this activity (2). In bacteria, lipoproteins are synthesized as precursors with a signal peptide linked to the amino terminus. The signal sequence is cleaved by signal peptidase II (36) during secretion, and the cysteine residue at the cleavage site is modified by addition of lipid. The activity of

the lipoprotein-specific signal peptidase II can be inhibited by globomycin.

Antibodies directed against a 40K outer membrane protein have been shown to prevent disease in a passive protection experiment (11), and a seroreactive protein of 40K is present in all *H. somnus* isolates tested (8). Therefore, we decided to investigate the LppA protein by using recombinant DNA technology.

In this paper, we report the cloning of the *lppA* gene in *E. coli*, determination of the nucleotide sequence, and initial characterization of the native, as well as the recombinant, LppA protein.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions. *E. coli* DH5 α [ϕ 80 *lacZ* Δ M15 *endA1* *recA1* *hsdR17* (r_K^- m_K^+) *supE44* *thi-1* λ^- *gyrA96* *relA1* Δ (*lacZYA-argF*)U169/F' *lacI*^q *proAB*⁺ *lacZ* Δ M15 Tn5 (Km^r)] and JM105 [*endA1* *thi* *rpsL* *sbcB15* *hsdR4* Δ (*lac-proAB*)/F'*traD36* *proAB*⁺ *lacI*^q Δ M15] are from the laboratory collection, and CC118 [*aroD139* Δ (*ara leu*)7697 Δ *lacX74* *phoA* Δ 20 *galE* *galK* *thi* *rpsE* *rpoB* *argE*(Am) *recA1*] is from C. Manoil. *E. coli* strains were grown in Luria broth (LB) or M63 (27). Ampicillin was used at 100 μ g/ml and kanamycin was used at 25 μ g/ml unless otherwise indicated. *H. somnus* HS25 has been used in earlier experiments to induce experimental haemophilosis in calves (14). Growth conditions for strain HS25, plasmid pGH433, and construction of the expression library have been described elsewhere (33).

DNA techniques. Restriction enzymes, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA ligase, and exonuclease III were used as recommended by the suppliers. DNA sequencing was accomplished by the chain termination method, essentially as described by Messing (26).

Primer extension mapping. RNA from exponentially growing cultures of HS25, JM105/pMS65, and JM105/pGH433 was isolated by a hot phenol method (35). A synthetic

* Corresponding author.

† Published with permission of the director as journal series no. 130.

‡ Present address: Statens Seruminstitut, Artillerivej 5, DK-2300, Copenhagen S, Denmark.

oligonucleotide complementary to the *lppA* mRNA was labelled at the 5' end with [³²P]phosphate and used as the specific primer in the primer extension experiments. The reverse transcription reaction was carried out in a 25- μ l reaction volume containing 10 μ g of RNA, 20 ng of labelled primer, 200 U of reverse transcriptase, 50 mM Tris-HCl, 75 mM KCl, 10 mM dithiothreitol, 3 mM MgCl₂, and each of the four deoxynucleoside triphosphates (each 500 μ M) (pH 8.3).

Transposon *TnphoA* mutagenesis. Fusions of *lppA* to *TnphoA* were created with λ ::*TnphoA* (12). In this system, alkaline phosphatase (AP) activity is obtained only if *TnphoA* transposes onto a DNA sequence in such a way that AP is fused in frame and downstream of an expressed coding sequence containing appropriate membrane insertional sequences (17, 23, 24). Plasmid pMS22 was transformed into strain CC118. The resulting strain was infected with λ ::*TnphoA* and grown for 15 h at 30°C. Aliquots were plated on LB agar supplemented with kanamycin (300 μ g/ml), ampicillin (100 μ g/ml), and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) (40 μ g/ml). The plates were incubated at 30°C for 2 to 3 days, and plasmid DNA was extracted from five pools of blue colonies and used to transform CC118 cells. Individual AP⁺ (blue) colonies were isolated at 37°C, and their plasmid DNA was analyzed by restriction mapping.

PAGE and immunoblotting. SDS-PAGE of *H. somnus* and *E. coli* proteins was performed in the Laemmli system (22) or by using tricine-SDS-polyacrylamide gels with a 16.5% T-6% C separating gel (29). Transfer of proteins onto a nitrocellulose membrane was performed as recommended by the manufacturer. Blots were incubated with bovine serum diluted 1:500 with Tris-buffered saline (TBS)-1% bovine serum albumin (BSA) (TBS is 10 mM Tris-Cl [pH 7.5], 140 mM NaCl) for 2 h. The antiserum used was bovine hyper-immune serum against live *H. somnus* HS25. After three washes in TBS containing 0.05% Tween 20, seroreactive proteins were detected with goat anti-bovine immunoglobulin G coupled to AP (Kirkegaard & Perry) at 1:5,000 in TBS-1% BSA. AP activity was visualized by using the nitroblue tetrazolium-BCIP system as described by the supplier (Promega). Prestained or nonstained protein standards were obtained from Bio-Rad.

Labelling of proteins with [³H]palmitate. *E. coli* DH5 α F'IQ harboring the specified plasmids was grown in M63 medium supplemented with glycerol (0.5%, wt/vol) and Casamino Acids (2%, wt/vol). *H. somnus* HS25 was grown in BHI-TT medium. To exponentially growing cells (4 \times 10⁸ cells per ml), [³H]palmitate (5 mCi/ml) was added to a final concentration of 50 μ Ci/ml, and incubation was continued for 2 h. Labelling was terminated by precipitation with trichloroacetic acid (10%, wt/vol) for 30 min on ice. When indicated, globomycin (10 mg/ml in dimethyl sulfoxide) was added at 100 μ g/ml 5 min prior to the addition of palmitate (globomycin was a gift from M. Arai, Sankyo Co., Tokyo, Japan). Proteins were pelleted by centrifugation at 15,000 \times g for 20 min, and the pellets were washed twice with methanol to remove lipids. The dried pellets were resuspended in sample buffer and analyzed by tricine-SDS-PAGE. The radiolabelled protein bands in the dried gel were detected by fluorography.

Oligonucleotide-directed mutagenesis. A 33-residue synthetic oligonucleotide with the sequence 5'-TGTATTATTAGCAGCTGGTAATGAAAAAATAA was synthesized to alter the Cys-22 residue of the LppA protein (the underlined base differs from the wild-type sequence). Oligonucleotide-directed mutagenesis was performed as described by Kunkel

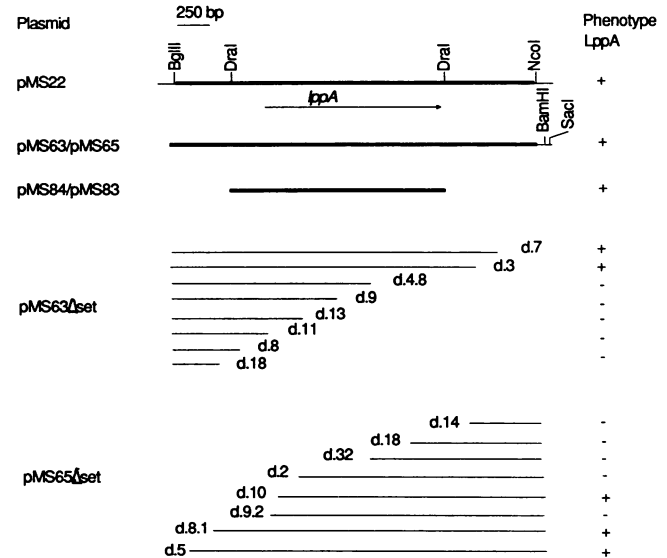


FIG. 1. Structures and properties of specific plasmids. The top line shows a partial restriction map of plasmid pMS22, with sites relevant to this discussion indicated. The arrow in the next line indicates the location and direction of transcription of *lppA*. The shaded bars in the two lines beneath the arrow illustrate the DNA cloned in each of the plasmids listed at the left. Where two names are separated by a slash, the fragment was cloned in both orientations. The lower two sets of lines indicate the DNA remaining in the deletion plasmids used for determining the DNA sequence. The ability of the plasmids to direct the synthesis of LppA in JM105 is shown at the right.

(21). The point mutation in the resulting plasmid pMS67 was verified by DNA sequencing.

RESULTS

Cloning of *lppA* in *E. coli*. A genomic library of *H. somnus* HS25 DNA was constructed by cloning 2- to 7-kb fragments, generated by partial *Sau3A* restriction, into the plasmid expression vector pGH433, and positive transformants were detected by the colony blot method (10) using antiserum against *H. somnus* HS25. Twenty-eight positive clones were identified and kept for further analysis. To identify the plasmid-encoded proteins reacting with the serum, whole cell lysates of isopropylthiogalactoside (IPTG)-induced cell cultures were examined by PAGE and subsequent Western immunoblotting. Three plasmids encoding a seroreactive protein with an M_r of approximately 40,000 were identified. One of these, with a DNA insert of 2 kb, was designated pMS22. Preimmune serum did not react with the seroreactive 40K protein (data not shown). By using the radiolabelled insert of pMS22 as a probe, it was shown that the three plasmids contained common sequences, indicating that the 40K recombinant proteins are identical (data not shown). Serum from calves immunized with the recombinant LppA protein reacts strongly with a 40K protein present in an outer membrane-enriched fraction of *H. somnus* (data not shown).

Analysis of recombinant plasmids. To subclone the *lppA* gene and construct plasmids suitable for exonuclease III degradation of the cloned region, the *BglIII-NcoI* fragment of pMS22 was cloned into pTZ18R (Fig. 1). Two plasmids, pMS63 and pMS65, with the insert in opposite orientations, were obtained. Both expressed the *lppA* protein, indicating that the gene is being transcribed from a promoter located on

```

AAAAATCCATTGATAGCAATCAGTTTTATCTGAAATTGGTACAAAAATAAATACTATTTTTAGTATGA 70
ATACCAGTGCAGAACTTTACGACTAGAACTTCGTTTACGCTCGCCGTTGATGCAGGGTTATTGGGGTG 140
TTCCTTAAATGCCCTTTGAAAATTACCAACTGAATGAAGCGTGGACTTGGGAAAAACAGGCTTTAGTTCGT 210
TGTAGGGCGGTATACGGCGATTTGATTTATGTGAACGCTTTGAAAAATTCGTTGTAATGTGCTTTCAG 280
CTCCAAGAAATGTGGAACAGCTGAAGCAAGATATACGAGAGATGCGTCAAAAAATGTATCATCATCTCTC 350
TAAACATAAAACGGACGAATTAATATTAAGACTGATTTGGGCGGTATCACAGATATTGAGTTTATTGCA 420
CAATACTTAGTTTTAGCTTATGCTCCCAACTAGCATTAAACACGTTGGTCTGATAATGTAGGATATTT 490
GACTGTATGGCTGAAAGTGCGGTGATTTCAACAAGAGTTCCACAAGTCAAAAAATGCTATGTAATTT 560
          DraI
TACGAAACCAAATTCATCATTAAATTTATTAGGTCAAGAACCATTATTAATGCACAACATTTTAGCAA 630
GGAAAGAACGTTTATTCTCAATACATGGAAAAGTTTATTGGAATGAATGAACCTATAATTGCCCTAAAAT 700
          -35          -10          +1
CAGCATATGATAAGAAATTATTTATCATTGTTATTTCTTTGTTATGCTATGCAGACCTTTAACTTACAT 770
TAACAAAATGAGAAATAAACGATGAAATTAATAAATCACTTTTGGTCGGCACATTAGTCGCCTCAACTGT 840
          SD          M K L N K S L L V G T L V A S T V
ATTATTAGCAGCTTGTAAATGAAAAAATAAAGCGGAAACAACGCCAACTGAACCGGTTACAGTTGCAGAA 910
          L L A A C N E K N K A E T T P T E P V T V A E
ACTCAAGCTCAACCTGACGTTCAAGGAAAAACTGAAACAACCTCATCTGAATCAACCGCAATTGAAAAATA 980
          T Q A Q P D V Q G K T E T T S S E S T A I E N
CACAATCTGATGCTCAAGAAAAACTGAGACAACCTCAGTTGAAACAACCTCGACTGAACCAACCGCAGC 1050
          T Q S D A Q E K T E T T S V E T T S T E P T A A
TGGAACACACAACCTGAATCTCAAGAAAAAGTTGTTTCAGAAAAAGTGAAGAGTTGTTCAAGAAATTT 1120
          G N T Q P E S Q E K V V S E K S E T V V Q E I
CTTAATCAGTTTAAACAATACAGTTACGATCCAATTTGGTGGGGTATCAGAGTGAAAAAATAGAGGGTGAAG 1190
          L N Q F N N T V T I Q L V G Y Q S E K I E G E
ATACTTTATCTTTCGTTTATAACGTTAAGAATAAAGGTGATAAAGCAATCAAAGAATTCAGTGGTATAA 1260
          D T L S F V Y N V K N K G D K A I K E L Q W Y N
CCTGTTTTCTTTAATTCGACTCTGGTAGAGCCTCTTCAATAGCCTATTCTTTTGGAGATACGCTTGCT 1330
          L V F F N S T L V E P L S I A Y S F E D T L A
          pho47
CCGGAAGGGCAGGGCGAAATAAAATTAACAAAATTAGCTAAAATTCATGCTGAAGAGATTCGTGCAGATA 1400
          P E G E G E I K L T K L A K T Y A E E I R A D
TACTAAAACCGGAAGCTAATCTCAATTTAGSCCAATAATTGCAGGTGCAATTTTTTTGAAGAGGGTAC 1470
          I L K P E A N L Q F T S P I I A G R I I F E D G T
          pho49
GCAATTAGTTGTAACACTACAGATGAAGAGCTTACTCAATCTTTACAGCAAATTTTAACGCAATAATTTTA 1540
          Q L V V T T D E E L T Q S L Q Q I L T Q END
          DraI
AAAATAATTATTCAACGCATTAGTTATCTATCCGCTCTTACAAATCTATAATATTTATAAATAACTACAA 1610
AAAGTTATCAATAAGATTTTATAGATTGGTAGATCGGTATGTTTCCGCATCGAAATCTACTGCCATT 1680
ATTGGCGAAACCGAAAGAAATTCGTCGTAAAAGCGTGCAGACCAACAGAAAAAGTGTGAAGAAAA 1750
AAAGCTGAGAAATTTGCTAAAAATCAGCTCAACAAACCGCACTTAATAATAAAAAATTTCTGCGAGAAATC 1820
ATGTAAAAAAATAACCCCTCTTAAACAAGAAGAGGGTGAATAATCAATTTACCATTGGTACCCTATAGA 1890
AACTGAACCTGCCATTTTGCCTTGAGAAATTTCTATTTCCTTGAATTTAAGCATAATCTTACGTTATCAC 1960
TCATACGAGAATAACCAATCGCCAT 1984
    
```

FIG. 2. Nucleotide sequence of the *H. somnus lppA* region. The sequence of the antisense strand is shown, with numbering starting from the 5' end. SD, Shine-Dalgarno sequence (30). The transcriptional start of the *lppA* gene is indicated by +1, and the positions of the two *lppA::phoA* fusions are shown.

the insert DNA. To generate a series of nested deletions, plasmids pMS63 and pMS65 were each cut at the unique *SacI* and *BamHI* sites (Fig. 1) and subjected to exonuclease degradation, removal of overhang by S1 nuclease, and religation. A number of plasmids were analyzed, and the extent of the degradation (as judged by restriction mapping or DNA sequencing) was compared with the phenotype (Fig. 1). It appears from this deletion experiment that the *lppA* gene is located between the deletion endpoints of d.3 and d.8.1 because plasmids with a larger insert are LppA⁺, whereas plasmids with deletions going further into the insert are LppA⁻. This is true with one exception, namely, d.10, which produces a seroreactive truncated version of the LppA protein with an *M_r* of approximately 37,000 (data not shown). DNA sequencing of the deletion endpoints of the two plasmids d.9 and d.10 revealed that in d.10, the α peptide of *lacZ* is fused in frame with the *lppA* open reading frame (see below), thereby allowing the gene to be transcribed from *lacP* or another vector-encoded promoter and translated from the *lacZ* translational start site. In contrast, *lacZ* in d.9 is fused out of frame with the *lppA* open reading frame.

DNA sequencing and analysis. The complete DNA sequence of both strands was determined by the dideoxy method with modified T7 DNA polymerase and single-stranded DNA as the template. The sequence is shown in Fig. 2. Only one open reading frame long enough to encode the *lppA* gene product is present on the sequenced DNA. It begins with an ATG codon located at positions 791 to 793 and terminates with the TAA stop codon at positions 1532 to 1534. This open reading frame would encode a polypeptide with a molecular weight of 27,072. The ATG start codon is preceded by the purine-rich sequence AATGAG (underlined bases are complementary to 16S rRNA), which serves as a ribosome binding site in *E. coli* (31).

The proposed reading frame has been confirmed by sequencing two independent *lppA::TnphoA* gene fusions (Fig. 2). Further proof that the indicated open reading frame is *lppA* was obtained by subcloning the *DraI* fragment of pMS22 (Fig. 1) into the *SmaI* site of pTZ18R and generating pMS83 and pMS84, with the insert in opposite orientations. *DraI* cuts 209 bp upstream of the putative ATG start codon and immediately downstream of the TAA stop codon. The LppA protein was expressed in JM105 harboring both plas-

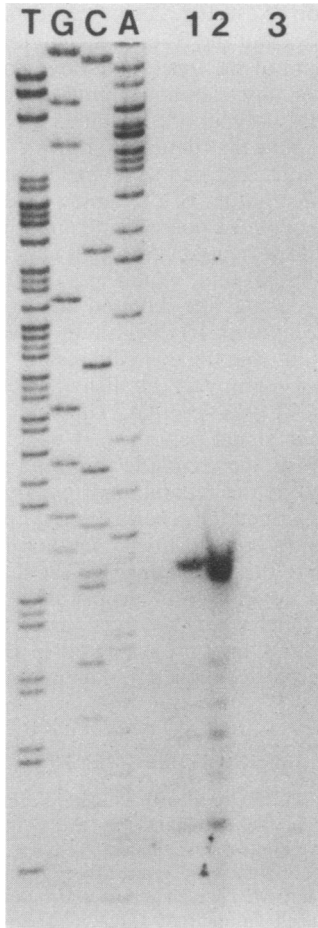


FIG. 3. Levels of *lppA* mRNA measured by the primer extension mapping method. An autoradiogram of a polyacrylamide sequencing gel used to analyze the primer extended fragments is depicted. The primer extension reactions were done as described in Materials and Methods. The cellular RNAs used were obtained from HS25 (lane 1), JM105/pMS65 (lane 2), or JM105/pGH433 (lane 3). Identical amounts of each primer extension reaction mixture were loaded onto the gel. The dideoxy sequencing ladder was generated by using the same primer, but unlabelled. The ladder is shown labelled as the complementary base above each lane.

mids (data not shown). The N-terminal part of the predicted polypeptide strongly resembles a signal peptide, and the amino acid sequence Leu-Leu-Ala-Ala-Cys at positions 842 to 856 is highly homologous to the consensus cleavage site found in bacterial lipoproteins (36).

Identification of the 5' terminus of *lppA* mRNA. The 5' terminus of the *lppA* transcript was determined by primer extension mapping. The DNA used as the primer was a synthetic 5'-end-labelled oligonucleotide complementary to nucleotides between 817 and 835. mRNA was isolated from *H. somnus* HS25 and from two *E. coli* strains, JM105/pMS65 (*LppA*⁺) and JM105/pGH433 (*LppA*⁻). The data presented in Fig. 3 show that one major *lppA* transcript initiated with the A residue at position 756 is produced in both HS25 and JM105/pMS65. No product was observed in cells harboring the plasmid vector pGH433 (Fig. 3, lane 3). A Pribnow box and -35 region, characteristic of *E. coli* promoters (15), are located at positions 744 through 749 (TATGCT) and positions 722 through 727 (TTATCA), respectively.

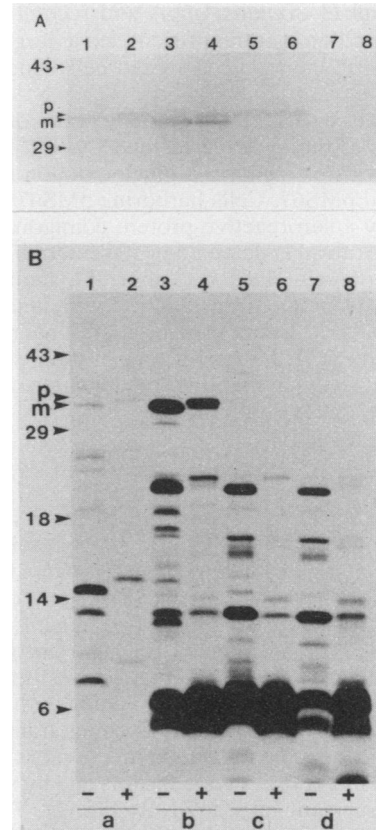


FIG. 4. Lipid modification and processing of the wild-type and mutated *LppA* proteins. Shown are an immunoblot (A) and tricine-SDS-PAGE fluorogram (B). Cells of *H. somnus* HS25 (a) and of *E. coli* DH5 α F'IQ harboring pMS65 (b), pMS67 (c), or the vector pTZ18R (d) were labelled with [³H]palmitate for 2 h in the absence (-) or presence (+) of globomycin. Whole cell proteins were precipitated at 10% (wt/vol) trichloroacetic acid. The equivalent of 0.2 ml of the cultures was applied to each lane. The locations of the mature (m) and precursor (p) forms of the *LppA* protein are shown. The mobilities of the molecular weight standard proteins are indicated in kilodaltons.

Posttranslational modification of the *LppA* protein. Because the deduced amino acid sequence of the *LppA* protein contains a sequence identical to the consensus sequence Leu-Ala(Gly)-Ala(Gly)-Cys for lipid modification in *E. coli* (36), the *lppA* gene product may be a lipoprotein. To test whether the *LppA* protein was lipid modified, [³H]palmitate was incorporated into *H. somnus* HS25 and two *E. coli* strains, DH5 α F'IQ/pMS65 and DH5 α F'IQ/pTZ18R. Proteins from whole cell lysates were separated by PAGE and transferred to a nitrocellulose membrane. The *lppA* gene product was identified by immunoblotting with antiserum against HS25. At least 10 *H. somnus* proteins are labelled with palmitate (Fig. 4B, lane 1). One of these is a 40K protein which reacts with *H. somnus* antiserum, showing that it is the *lppA* gene product (Fig. 4A, lane 1). Palmitate was also incorporated into the recombinant *lppA* gene product, since a radiolabelled, immunoreactive 40K protein comigrating with the *LppA* protein from HS25 was detected in cells harboring pMS65 but not in the plasmid vector pTZ18R (Fig. 4, lanes 3 and 7). Thus, the *H. somnus lppA* gene product is lipid modified in *E. coli*. Treatment of cells with globomycin leads to the accumulation of unprocessed lipoprotein, and

both the natural *H. somnus* LppA and recombinant LppA proteins are predominantly present as a larger, putative precursor form in globomycin-treated cells (Fig. 4, lanes 2 and 4).

To determine whether lipid modification of the LppA protein occurs at the cysteine residue Cys-22, the cysteine codon (TGT) was changed to a glycine codon (GGT), generating plasmid pMS67. Cells harboring pMS67 are LppA⁺. However, only a seroreactive protein comigrating with the larger precursor form is detected in a Western blot (Fig. 4A, lane 5). Globomycin does not alter the mobility of the mutated LppA protein (compare lane 5 with lane 6), indicating that the mutated LppA protein is no longer a substrate for signal peptidase II. Moreover, this protein is not labelled with palmitate, showing that lipid modification occurs at the Cys-22 residue.

DISCUSSION

The major antigen of *H. somnus* HS25 is a 40K outer membrane protein. In this report, this protein is named LppA and the corresponding gene is named *lppA*. We present data which show that the *lppA* gene of HS25 has been isolated on a 2-kb *Sau3AI* fragment in plasmid pGH433. Nucleotide sequencing revealed an open reading frame starting at an ATG codon (positions 791 to 793), which encodes a polypeptide of 247 amino acids with a predicted molecular weight of 27,072. The reading frame has been confirmed by sequencing the fusion junction of two independently isolated *lppA::phoA* gene fusions and by subcloning a 957-bp *DraI* fragment into pTZ18R (Fig. 1). The resulting plasmid, pMS84, expressed the full-length *lppA* gene product, showing that although the deduced polypeptide has a predicted molecular weight of at least 13,000 less than the one observed, this frame does encode the LppA protein. We cannot yet explain the discrepancy between the deduced and the observed molecular weights; however, we note that treatments such as prolonged boiling under reducing conditions, boiling at low pH, incubation at 37°C without β-mercaptoethanol, denaturation with urea, and subsequent electrophoresis in the presence of urea did not alter the mobility of the LppA protein in SDS-PAGE (data not shown). Moreover, anomalous migration in SDS-PAGE is characteristic of many lipoproteins. For example, OsmB migrates as a 10K polypeptide although it has an *M_r* of 4,230 (20), the 36K lipoprotein of *E. coli* had an *M_r* of 54,000 on gels (32), and peptidoglycan-associated lipoprotein is a 16K polypeptide but has an *M_r* of 20,000 (6).

The *lppA* gene showed a biased codon usage, with 82% (195 of 247) of the codons having either A or T at the third position. This bias is consistent with that observed for other *H. somnus* coding regions (33). Also, we did not find any significant homology between the LppA amino acid sequence and sequences compiled in GenBank release 65, indicating that the *H. somnus lppA* gene product is a novel protein.

The results of the 5'-end mapping of *lppA* in *H. somnus* and of the cloned gene in *E. coli* indicate that the *H. somnus* promoter is active in *E. coli*. This is not surprising since the proposed -10 and -35 regions (Fig. 2) conform to the -10 and -35 regions of the *E. coli* consensus promoter (15). The transcriptional termination site of the *lppA* gene has not been established. However, a Northern (RNA) blot of *H. somnus* RNA separated on a denaturing agarose gel indicated that LppA is specified by one single transcript approximately 1,300 nucleotides long (data not shown). With the start point

located at position 757 in Fig. 2, this suggests that transcription terminates beyond the 3' end of the cloned DNA. The region downstream of the *lppA* gene does not contain open reading frames of any significant length. Also, the LppA protein is the only polypeptide specified by the *H. somnus* insert in *E. coli* minicells (data not shown). Therefore, it is likely that *lppA* is transcribed as a single cistron.

The hydrophobic nature of the amino acids at the N-terminal part of the deduced polypeptide and the presence of a lipoprotein box, Leu-Leu-Ala-Ala-Cys, suggested that *lppA* specified a new lipoprotein, which was confirmed by direct experiments. By using site-directed mutagenesis and by labelling of proteins with [³H]palmitate, it was shown that LppA is lipid modified at the cysteine residue in the cleavage region (Fig. 4). Globomycin abolished processing of the precursor of the LppA protein, showing that it is the lipoprotein-specific signal peptidase II which catalyzes the cleavage. Therefore, we conclude that lipid modification of the native as well as the recombinant *lppA* gene product is similar to that described for other bacterial lipoproteins. It is also apparent that *H. somnus* must contain a signal peptidase II closely related to the *E. coli* analog. The data presented in this report add a novel protein, LppA, to a growing list of bacterial lipoproteins which are recognized as strong antigens. These proteins are often located in the outer membrane and are therefore exposed to the immune system of the host.

ACKNOWLEDGMENTS

We thank M. Parker for help with the site-directed mutagenesis, G. Gerlach for valuable comments on the manuscript, and N. Rawlyk for expert technical assistance.

This work was supported by grants from the Danish National Science Foundation to M. Theisen and from the Natural Science and Engineering Research Council of Canada, The Garfield Weston Foundation, and the Canadian Bacterial Disease Network (NSERC Center of Excellence Program).

REFERENCES

- Anderson, B. E., B. R. Baumstark, and W. J. Bellini. 1988. Expression of the gene encoding the 17-kilodalton antigen from *Rickettsia rickettsii*: transcription and posttranslational modification. *J. Bacteriol.* **170**:4493-4500.
- Bessler, W. G., M. Cox, A. Lex, B. Suhr, K.-H. Wiesmuller, and G. Jung. 1985. Synthetic lipopeptide analogs of bacterial lipoprotein are potent polyclonal activators for murine B lymphocytes. *J. Immunol.* **135**:1900-1905.
- Bessler, W. G., and B. P. Ottenbreit. 1977. Studies on the mitogenic principle of the lipoprotein from the outer membrane of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **76**:239-246.
- Bessler, W., K. Resch, E. Hancock, and K. Hantke. 1977. Induction of lymphocyte proliferation and membrane changes by lipopeptide derivatives of the lipoprotein from the outer membrane of *Escherichia coli*. *Z. Immun.* **153**:11-22.
- Bricker, T. M., M. J. Boyer, J. Keith, R. Watson-McKown, and K. S. Wise. 1988. Association of lipids with integral membrane surface proteins of *Mycoplasma hyorhinis*. *Infect. Immun.* **56**:295-301.
- Chen, R., and U. Henning. 1987. Nucleotide sequence of the gene for peptidoglycan-associated lipoprotein of *E. coli* K-12. *Eur. J. Biochem.* **163**:73-77.
- Corbeil, L. B. 1990. Molecular aspects of some virulence factors of *Haemophilus somnus*. *Can. J. Vet. Res.* **54**:S57-S62.
- Corbeil, L. B., J. E. Arthur, P. R. Widders, J. W. Smith, and A. F. Barbet. 1987. Antigenic specificity of convalescent serum from cattle with *Haemophilus somnus*-induced experimental abortion. *Infect. Immun.* **55**:1381-1386.
- Dev, I. K., R. J. Harvey, and P. H. Ray. 1985. Inhibition of

- prolipoprotein signal peptidase by globomycin. *J. Biol. Chem.* **260**:5891-5894.
10. French, B. T., H. M. Maul, and G. G. Maul. 1986. Screening cDNA expression libraries with monoclonal and polyclonal antibodies using an amplified biotin-avidin-peroxidase technique. *Anal. Biochem.* **156**:417-423.
 11. Gogolewski, R. P., S. A. Kania, H. D. Liggitt, and L. B. Corbeil. 1988. Protective ability of antibodies against 78- and 40-kilodalton outer membrane antigens of *Haemophilus somnus*. *Infect. Immun.* **56**:2307-2316.
 12. Gutierrez, C., J. Barondess, C. Manoil, and J. Beckwith. 1987. The use of the transposon *TnphoA* to detect genes for cell envelope proteins subject to a common regulatory stimulus: the analysis of osmo-regulated genes in *E. coli*. *J. Mol. Biol.* **195**:289-297.
 13. Hanson, M. S., and E. J. Hansen. 1991. Molecular cloning, partial purification, and characterization of a haemin-binding lipoprotein from *Haemophilus influenzae* type b. *Mol. Microbiol.* **5**:267-278.
 14. Harland, R. J., A. A. Potter, and J. C. L. Schuh. 1990. Development of an intravenous challenge model for *Haemophilus somnus* disease in beef calves. *Conf. Res. Work. Anim. Dis.* **71st** 29:6.
 15. Harley, C. B., and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**:2343-2361.
 16. Harris, F. W., and E. D. Janzen. 1990. The *Haemophilus somnus* disease complex (hemophilosis): a review. *Can. Vet. J.* **30**:816-822.
 17. Hoffman, C. S., and A. Wright. 1985. Fusion of secreted proteins to alkaline phosphatase: an approach for studying protein secretion. *Proc. Natl. Acad. Sci. USA* **82**:5107-5111.
 18. Hubbard, C. L., F. C. Gherardini, P. J. Bassford, and L. V. Stamm. 1991. Molecular cloning and characterization of a 35.5-kilodalton lipoprotein of *Treponema pallidum*. *Infect. Immun.* **59**:1521-1528.
 19. Humphrey, J. D., and L. R. Stephens. 1983. *Haemophilus somnus*: a review. *Vet. Bull.* **53**:987-1004.
 20. Jung, J. U., C. Gutierrez, and M. R. Villarejo. 1989. Sequence of an osmotically inducible lipoprotein gene. *J. Bacteriol.* **171**:511-520.
 21. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* **82**:488-492.
 22. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
 23. Manoil, C., and J. Beckwith. 1985. *TnphoA*: a transposon probe for protein export signals. *Proc. Natl. Acad. Sci. USA* **82**:8129-8133.
 24. Manoil, C., and J. Beckwith. 1986. A genetic approach to analyzing membrane protein topology. *Science* **233**:1403-1408.
 25. Melchers, F., V. Braun, and C. Galanos. 1975. The lipoprotein of the outer membrane of *Escherichia coli*: a B lymphocyte mitogen. *J. Exp. Med.* **142**:473-482.
 26. Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* **101**:21-78.
 27. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 28. Nelson, M. B., M. A. Apicella, T. F. Murphy, H. Vankeulen, L. D. Spotila, and D. Rekosh. 1988. Cloning and sequencing of *Haemophilus influenzae* outer membrane protein P6. *Infect. Immun.* **56**:128-134.
 29. Schagger, H., and G. von Jagow. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**:368-379.
 30. Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to non-sense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71**:1342-1346.
 31. Stormo, G. D., T. D. Schneider, and L. M. Gold. 1982. Characterization of translational initiation sites in *Escherichia coli*. *Nucleic Acids Res.* **10**:2971-2996.
 32. Takase, I., and M. Matsuhashi. 1987. Genes encoding two lipoproteins in the *leuS-dacA* region of the *Escherichia coli* chromosome. *J. Bacteriol.* **169**:5692-5699.
 33. Theisen, M., and A. A. Potter. 1992. Cloning, sequencing, expression, and functional studies of a 15,000-molecular-weight *Haemophilus somnus* antigen similar to *Escherichia coli* ribosomal protein S9. *J. Bacteriol.* **174**:17-23.
 34. Thirkell, D., A. D. Myles, and W. C. Russell. 1991. Palmitoylated proteins in *Ureaplasma urealyticum*. *Infect. Immun.* **59**:781-784.
 35. von Gabin, A., J. G. Belasco, J. L. Schottel, and A. C. Y. Chang. 1983. Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. *Proc. Natl. Acad. Sci. USA* **80**:653-657.
 36. Wu, H. C., and M. Tokunaga. 1986. Biogenesis of lipoprotein in bacteria. *Curr. Top. Microbiol. Immunol.* **125**:127-157.