INFECTION AND IMMUNITY, July 1994, p. 2792–2799
0019-9567/94/$04.00+0

Vol. 62, No. 7

A Family of Surface-Exposed Proteins of 20 Kilodaltons in the Genus *Borrelia*

CAROL J. CARTER,[1] SVEN BERGSTRÖM,[2] STEVEN J. NORRIS,[3] AND ALAN G. BARBOUR[1]*

Departments of Microbiology and Medicine, University of Texas Health Science Center, San Antonio, Texas 78284-7758[1]; Department of Microbiology, University of Umeå, S-901 87 Umeå, Sweden[2]; and Department of Pathology and Laboratory Medicine, University of Texas Medical School, Houston, Texas 77225[3]

Received 30 November 1993/Returned for modification 23 December 1993/Accepted 1 March 1994

Relapsing fever and Lyme disease spirochetes of the genus *Borrelia* display at their surfaces abundant lipoproteins: Vmp proteins in *Borrelia hermsii* and Osp proteins in *Borrelia burgdorferi*. Vmp and Osp proteins largely determine serotype specificity, and neutralizing antibodies of infected or immunized animals are directed at them. For the present study, we examined *B. hermsii* serotype 33, which is unique among strain HS1 serotypes in the low frequency of switches to other serotypes during infections and in vitro cultivation. Failing to clone the complete *vmp33* gene, we accomplished its further characterization by (i) determining three partial amino acid sequences, (ii) designing oligonucleotide primers based on these amino acid sequences, (iii) cloning and sequencing the central portion of *vmp33*, and (iv) using outwardly directed primers and the inverse PCR to clone the 5′ and 3′ ends of the gene and flanking regions. The transcriptional start site was identified by primer extension analysis. Vmp33 was a polypeptide of 211 amino acids; the three partial amino acid sequences were identified in the open reading frame. Vmp33 was found to be more similar to other 20-kDa Vmp proteins of *B. hermsii* and to OspC proteins of *B. burgdorferi* than it was to 35- to 39-kDa Vmp proteins of the same strain. Moreover, OspC proteins were more similar to Vmp33 than they were to OspA, -B, or -D proteins of *B. burgdorferi*. These sequence similarities were consistent with Western blot (immunoblot) findings of cross-reactions between Vmp33 and OspC with anti-Vmp33 and anti-OspC sera. The promoter for the expressed *vmp33* gene was found to be different from the expression site for other active *vmp* genes characterized to date. These results indicate that Vmp33 and other small Vmp's belong with OspC to a genus-wide family of 20-kDa proteins and that expression of these proteins may be coordinated with expression of other Vmp and Osp proteins in *Borrelia* spp.

The agents of relapsing fever and Lyme disease are spirochetes of the genus *Borrelia*. All members of this genus are host-associated parasites that cycle between vertebrate and arthropod hosts. In their hosts, these bacteria occupy various niches, including blood or hemolymph, neural tissue, and salivary glands (for a review, see reference (6). Efficient and frequent transfer between vertebrate and invertebrate hosts is essential for sustained maintenance of borrelias in nature. Although the features of Lyme disease and relapsing fever differ in several respects, their agents have this requirement in common. Consequently, it is conceivable that relapsing fever and Lyme disease borrelias have similar pathogenetic mechanisms to achieve their aims.

Phylogenetic studies of the genus *Borrelia* have shown that the Lyme disease group of species can be further distinguished from the cluster of species that cause relapsing fever in North America at the level of DNA sequences for rRNA and flagellin genes (23, 26, 29). There has been less comparative analysis of genes encoding surface proteins, which presumably would be under greater selective pressure from a host environment rich with antibodies and other immune effectors. If relapsing fever and Lyme disease agents have similar invasion mechanisms, a promising place to search for homologous virulence determinants is the cell surface. The sequences of four outer membrane lipoproteins, OspA, OspB, OspC, and OspD, of *Borrelia burgdorferi* and six alleles for the variable major protein, Vmp, of *Borrelia hermsii* have been determined (11, 14, 16, 20, 27, 31, 39). The Vmp proteins are found in two size ranges: large, between 37 and 40 kDa, and small, between 19 and 22 kDa (3, 7, 31). Both the large and small Vmp proteins are surface exposed (3, 4, 8, 9). Whereas OspA, -B, and -D proteins are between 29 and 35 kDa, the OspC proteins are within the size range of the small Vmp proteins. OspC, like OspA, -B, and -D, is surface exposed (39). Wilske et al. found evidence of antigenic cross-reactivity between the OspC protein of *B. burgdorferi* and a small Vmp of *B. hermsii* (39). For the present study, we cloned and determined the sequence for the Vmp gene of *B. hermsii* that was antigenically similar to OspC. In an analysis of this and other sequences of Osp and Vmp genes, we show that some Vmp proteins are members with OspC of a family of outer membrane proteins of the genus *Borrelia*.

* Corresponding author. Phone: (210) 567-3938. Fax: (210) 567-6612. Electronic mail address: alan@taliesin.uthscsa.edu.

MATERIALS AND METHODS

Strains and culture conditions. The origins of serotypes 3, 7, 17, and 24 of type strain HS1 of *B. hermsii* (ATCC 35209) have been described (7, 31). Serotype 33 of strain HS1 was previously designated serotype C (4, 9, 25); the population of serotype 33 used in these studies was started from a single cell (9). Isogenic variants B311 and B314 of strain B31 (ATCC 35210) of *B. burgdorferi* differ in their expression of OspA, -B, and -C; B311 is OspA$^+$ OspB$^+$ OspC$^-$, and B314 is OspA$^-$ OspB$^-$ OspC$^+$ (34). All borrelias were grown in BSK II broth and harvested by the methods described previously (2). Cells were counted in a Petroff-Hauser chamber by phase-contrast microscopy. *Escherichia coli* JM109 (40) bearing plasmid pUC8 or recombinant plasmid pUC8-B31/ospCL$^+$ (39), which expressed the OspC protein of strain B31 of *B. burgdorferi*, was

grown in L broth supplemented with carbenicillin. For induction of expression, *E. coli* cells were grown in the presence of 1 mM isopropyl-β-D-thiogalactopyranoside for 2 h.

Antibodies. The origin of the anti-Vmp33-specific monoclonal antibody H4825 has been given (8); ascitic fluid was used for this study. Polyclonal rabbit antiserum to purified OspC of *B. burgdorferi* (35) was kindly provided by Tom Schwan, Rocky Mountain Laboratories. The origin of the polyclonal rabbit antiserum to purified Vmp33 (formerly VmpC) was described previously (4).

Globomycin treatment, [^{3}H]palmitate labeling, PAGE, and Western blot (immunoblot) analysis. In a preliminary study, the MIC of the antibiotic globomycin (a gift of Sankyo Co. Ltd., Tokyo, Japan) for *B. hermsii* HS1 was 80 μg/ml. Globomycin was added to a culture of borrelias in BSK II medium at a cell density of 5×10^7/ml to a final concentration of 100 μg/ml. After 14 additional h of incubation, the cells were harvested as described previously (2) and washed three times with an equal volume of RPMI 1640 tissue culture medium without supplements (Mediatech). As a control, another culture was handled in the same manner but was not treated with globomycin. Borrelias were labeled with [9,10-^{3}H]palmitic acid (Dupont, Boston, Mass.) in BSK II medium; 2.5 mCi was added to a culture of 300 ml at a cell density of 10^6/ml. When the cell density reached 10^8/ml, the borrelias were harvested as described previously (2) and washed three times with 20 ml of RPMI 1640. Whole-cell lysates were subjected to polyacrylamide gel electrophoresis (PAGE) with 15% acrylamide as described previously (13). For Western blot analysis, proteins were transferred to nitrocellulose membranes, which were then blocked with 3% (wt/vol) dried nonfat milk in 10 mM Tris (pH 7.4)–150 mM NaCl (milk/TS) for 2 h (32). After the membranes were washed with milk/TS, they were incubated with ascitic fluid diluted 1:100 or rabbit antiserum diluted 1:500 in milk/TS. Alkaline phosphatase-conjugated recombinant protein A/G (Immunopure; Pierce Chemical Co., Rockford, Ill.) served as the second ligand. The blots were developed by using nitroblue tetrazolium chloride/5-bromo-4-chloro-3-indolyl-phosphatase *p*-toluidine salt (Pierce). Gels with radiolabeled proteins were processed for fluorography as described previously (14). Dried blots and developed films of fluorographs were scanned as gray-scale images at 300 dots per inch by using a flatbed image scanner (Microtek, Torrance, Calif.). The digitized images were saved as tag image file format files (8 bits per pixel), labeled with Adobe Illustrator for NeXT (Adobe Systems, Mountain View, Calif.), and printed on a 400-dots-per-inch laser printer (NeXT Computer, Redwood City, Calif.).

Protein extraction and partial amino acid sequencing. A fraction enriched for Vmp33 was obtained by using differential solubilization in the detergent octyl-glucopyranoside as described previously for Vmp7 and Vmp21 of *B. hermsii* (10). The final pellet was dissolved in trifluoroacetic acid, divided into aliquots, and then dried under vacuum. The Vmp33 fraction was treated either with endoproteinase Glu-C (Worthington) as described previously (10) or, after carboxymethylation, with endoproteinase Lys-C (Boehringer-Mannheim, Indianapolis, Ind.) (24). The resultant mixtures of peptides were separated by PAGE and then transferred to polyvinyl difluoride membranes (Millipore Corp., Bedford, Mass.) (24, 27). N-terminal amino acid sequences were determined by Per-Ingvar Ohlsson of the University of Umeå and Richard Cook of the Baylor College of Medicine by using model 477A sequenators (Applied Biosystems, Foster City, Calif.).

DNA procedures. Plasmid-enriched DNA was isolated from serotype 33 cells by the method of Hinnebusch and Barbour (19). The total RNA of borrelias was obtained as described by Meier et al. (25). Restriction enzymes from Boehringer-Mannheim were used as described in the manufacturer's recommendations. Probe DNA was labeled with [α-^{32}P]dATP either by nick translation with a commercial kit (Bethesda Research Laboratories, Gaithersburg, Md.) or with a Random Prime kit (Boehringer-Mannheim). Oligonucleotide probes were radiolabeled with [γ-^{32}P]dATP and T4 polynucleotide kinase and purified by passage through Nensorb 20 cartridges (DuPont). Hybridization of radiolabeled probes with the membranes was in a mixture containing 0.1% sodium dodecyl sulfate (SDS), 10× Denhardt's solution, 4× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), and 0.1 mg of denatured salmon sperm DNA per ml; membranes were washed in 6× SSC–0.1% SDS under high-stringency conditions (23). Custom oligonucleotide primers for sequencing and PCR were synthesized on an Applied Biosystems DNA synthesizer.

PCR. Enzymatic amplifications of the plasmid-enriched DNA were performed in 100-μl reaction volumes containing 1.25 U of *Taq* DNA polymerase (Perkin-Elmer, Norwalk, Conn.) and 50 pmol of each primer in a mixture containing 200 μM each deoxynucleoside triphosphate, 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, and 0.001% gelatin. Reactions were carried out in a thermal cycler (Perkin-Elmer) with an initial denaturation step of 94°C for 1.5 min and then 40 repetitions of 94°C for 1 min, 46°C for 1 min, and 72°C for 3 min.

Inverse PCR was performed essentially as described by Ochman et al. (28). DNA at a concentration of 100 μg/ml was digested with a restriction enzyme. After another extraction of the DNA to remove the enzyme, the DNA was diluted 1:100 and the free ends were ligated with T4 DNA ligase at 15°C for 14 h. The ligation products were precipitated twice with ethanol, resuspended in 20 μl of water, and denatured at 94°C for 1 min. PCR was carried out for 40 cycles under these conditions: 94°C for 30 s, 56°C for 30 s, and 70°C for 2 min. The amplified product was ligated to the plasmid vector pCRII and transformed into EC Sure cells (Stratagene, La Jolla, Calif.). Plasmid DNA was extracted by using Qiagen Maxi columns (Qiagen, Chatsworth, Calif.).

DNA sequence analysis. The DNA sequences of inserts of recombinant plasmids were determined in both strands by double-stranded dideoxy sequencing with Sequenase version 2.0 (U.S. Biochemicals, Cleveland, Ohio) with standard M13 forward and reverse primers or custom synthesized primers (21, 27). The transcriptional start site was identified by primer extension analysis of *B. hermsii* mRNA (5, 27). Sequences were analyzed by using the Genetics Computer Group set of programs, specifically GAP for pairwise alignments of sequences, PILEUP for multiple alignment and estimation of genetic divergence between sequences (15), and PEPPLOT for prediction of regions of hydrophilicity and hydrophobicity by the Kyte-Doolittle algorithm (18). The default parameters were used for these analyses. The BLAST program of Altschul et al. (1) was used for data base searches on Sun 690 MP and Cray Y-MP computers running the GenTools Release 3.0 suite of programs for the X Window system (Center for High Performance Computing of the University of Texas, Austin). The sequence of the *vmp33* gene has the GenBank accession number L24911.

RESULTS

Characterization of Vmp33. The Vmp7 and Vmp21 proteins of *B. hermsii* and OspA, -B, and -D proteins of *B. burgdorferi*

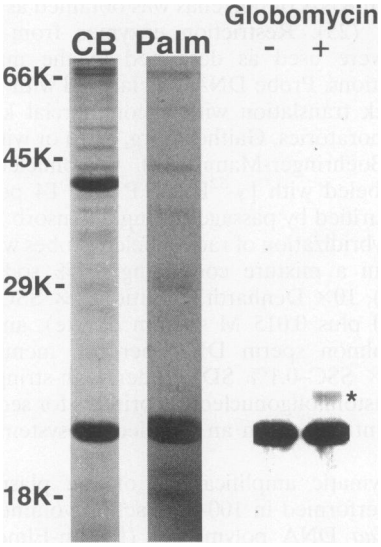

FIG. 1. Analysis of the Vmp33 protein. *B. hermsii* serotype 33 cells were grown in broth medium with (Palm) or without (CB) [^{3}H]palmitate and with (+) or without (−) globomycin. Whole-cell lysates of the cells were subjected to PAGE. The gels were stained with Coomassie blue (CB), prepared for fluorography (Palm), or transferred for Western blot analysis with the Vmp33-specific monoclonal antibody H4825 (Globomycin). An asterisk indicates an antibody-reactive protein in the globomycin-treated cells. The relative migrations of molecular weight standards (K = 1,000) in the gels are shown on the left.

had been identified previously as lipoproteins (12, 14, 27). To determine whether the Vmp33 was also a lipoprotein, borrelia cultures were radiolabeled with [^{3}H]palmitate or treated with the signal peptidase II inhibitor globomycin, as was done in previous studies (12, 14, 27). The leftmost two lanes in Fig. 1 show the Coomassie blue-stained proteins (CB) of serotype 33 in the gel and the [^{3}H]palmitate-labeled proteins (Palm) of this serotype in the fluorograph. The rightmost two lanes of the figure show the Western blot of serotype 33 borrelias with or without exposure to globomycin; the Vmp33-specific monoclonal antibody H4825 identified the location of Vmp33 in the gel and blots. Vmp33 was an abundant protein in the Coomassie blue-stained gel and had an estimated size of 20,000; the fluorograph demonstrated that it was the only protein heavily labeled with palmitate. There was no detectable labeling of the other abundant protein, which had an estimated size of 41,000, in the whole-cell lysate. Cells treated with globomycin, but not those without this exposure, contained a component of an estimated size of 22,000 that was recognized by the Vmp33-specific antibody. This would be expected if globomycin inhibited cleavage of a signal peptide from the *vmp33* gene product. These studies indicated that Vmp33 was a lipoprotein.

Antigenic similarity between the small Vmp proteins of *B. hermsii* and the OspC protein of *B. burgdorferi* was investigated with polyclonal antibodies to purified Vmp33 and OspC. Antibodies in polyclonal antiserum to OspC bound to Vmp33 as well as to native OspC expressed by isolate B314 and recombinant OspC expressed by *E. coli* (Fig. 2). The location of Vmp33 in the blots is indicated by monoclonal antibody H4825, which is specific for Vmp33 (8, 33). The monoclonal antibody did not bind to OspC (data not shown). The anti-OspC antibodies also bound weakly to Vmp3 (lane Bh3); they did not bind to the Vmp's of serotypes 7 and 17. Antibodies in polyclonal antiserum to purified Vmp33 bound to Vmp3 as well as to Vmp33 (Fig. 2). It had been shown previously that antibodies in this antiserum did not bind to the Vmp's of serotypes 7 and 21 (9). Antibodies to Vmp33 also bound to the native OspC protein of isolate B314 of *B. burgdorferi* and just detectably (Fig. 2, asterisk) to recombinant OspC protein. These findings confirmed an antigenic relatedness between OspC and Vmp33 (39) and further indicated that at least one other Vmp, namely Vmp3, was also similar to OspC.

Determination of the sequence of *vmp33*. All attempts to clone the entire gene for Vmp33 in either expression or genomic vectors were unsuccessful. These included approaches that had been successfully used in our laboratories to clone other Vmp genes and various Osp proteins of *B. burgdorferi*. Concluding that the presence of the complete gene, even in a nonexpressed state, was lethal or deleterious to an *E. coli* host, we set out to isolate DNA fragments representing internal sequences of the gene. For this, partial amino acid sequences were determined first. A fraction enriched for Vmp33 was obtained by using octyl-glucopyranoside and heat precipitation (10). Sequences could not be obtained from isolated complete protein, an indication that the end was acylated like other Vmp proteins (14). Accordingly, the protein was proteolyzed, and N-terminal sequences of isolated peptides were obtained. Three distinct and nonoverlapping sequences were obtained: (i) TVLDLSKISANIKNAVTFA; (ii) LNTAIDELLTA; and (iii) AQHNNLGQSAEAP. The first sequence was highly similar to the deduced sequence for Vmp3 of *B. hermsii*, beginning at residue 37 of the unprocessed polypeptide (TV LDLSKISANIKNASDFA) (31); the sequences differed at only 2 of 19 positions over this length. The second sequence resembled a peptide beginning at residue 180 of deduced Vmp3 protein (LNTAIDELLKA). The two aligned peptides differed in only 1 of 11 positions. There was little apparent similarity between the third peptide of Vmp33 and other Vmp or Osp proteins.

We concluded that the first peptide represented the N-terminal region of Vmp33 and the second peptide represented the C-terminal region. The third peptide could not be placed. On the basis of the amino acid sequences of peptides 1 and 2 and the codon usage of other *Borrelia* sp. genes (11, 14, 16, 17, 20, 27, 31, 39), the following inwardly directed oligonucleotide primers were synthesized: 5′-GCA AAT ATA AAA AAT GCG GTT-3′ (sense) and 5′-TGT TAA CAA TTC ATC AAT TGC-3′ (antisense). With these primers in a PCR with serotype 33 DNA, a 413-bp-long product was obtained. When this product was labeled and used in a probe for a Northern (RNA) blot of RNA of serotype 33 and serotype 7, the only RNA with a hybridizing band was that of serotype 33 (data not shown). This indicated that the amplified product was specific for the expressed *vmp* gene of serotype 33 (25).

The PCR product was ligated into a plasmid vector, which was transformed into *E. coli*. The inserts of three clones were sequenced. When used as a probe in a Southern blot of serotype 33 DNA, the cloned insert hybridized to a single 1.7-kb *Eco*RV restriction fragment. Inasmuch as there was not an *Eco*RV site in the sequenced DNA, and we lacked only an estimated 100 to 150 nucleotides on either side of the cloned insert, it was likely that the *Eco*RV fragment contained the 5′ and 3′ ends of the putative *vmp33* gene. Accordingly, a diluted *Eco*RV digest of serotype 33 DNA was self-ligated, and the products were used as the template DNA for inverse PCR. The outwardly directed primers represented sequences near the 5′ and 3′ ends of the cloned 413-nucleotide insert: 5′-TCT ACT TCT TGA ACA CTT GCA GC-3′ (sense) and 5′-CTG AAG AAC TTG GTA AGT TAA ATA C-3′ (antisense). Only the expected 1.3-kb PCR amplification product was obtained. It

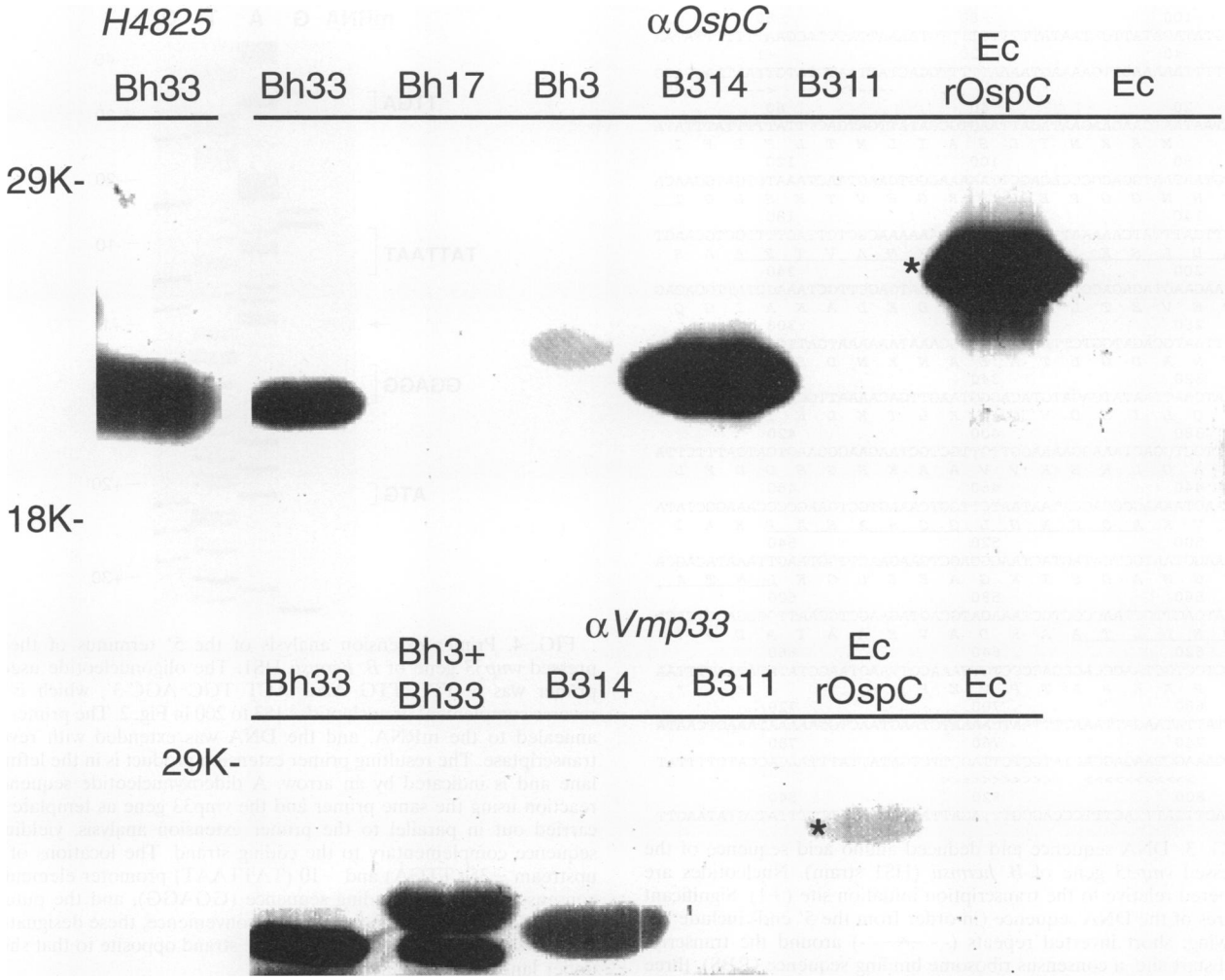

FIG. 2. Western blot analysis of native and recombinant proteins of *B. hermsii* and *B. burgdorferi* (see text for methods). (Top) Rabbit polyclonal antiserum to OspC (α*OspC*) of *B. burgdorferi* or monoclonal antibody to Vmp33 of *B. hermsii* (*H4825*) was incubated with blot membrane with the following samples: *B. hermsii* serotype 33 (Bh33), serotype 17 (Bh17), serotype 3 (Bh3), *B. burgdorferi* B314, *B. burgdorferi* B311, *E. coli* expressing recombinant OspC (Ec rOspC), and host *E. coli* (Ec). B311 is $OspA^+$ $OspB^+$ $OspC^-$, and B314 is $OspA^-$ $OspB^-$ $OspC^+$ (34). (Bottom) Some of the same samples, including equal amounts of Bh3 and Bh33 DNA in one lane (Bh3 + Bh33), were subjected to Western blot analysis with rabbit polyclonal antibody to Vmp33 (α*Vmp33*). Relative migrations of molecular weight standards (K = 1,000) in the gels are shown on the left.

was ligated to the plasmid vector, and the recombinant construct was transformed into *E. coli*. Three clones were sequenced over both strands by using custom primers that were based on data obtained on the first cloned insert. When this sequence was added to the first sequence, the sequence shown in Fig. 3 was obtained. An open reading frame began at nucleotide position +21 and ended at position 650; it would encode a polypeptide of 211 amino acids. The identity of this deduced protein with Vmp33 was confirmed by locating the three partial amino acid sequences within the open reading frame (Fig. 3). A consensus signal peptidase II site (38), FISC, was present at residues 16 to 19 of the open reading frame. The finding of a blocked N terminus in the attempt at sequencing the intact protein was consistent with acylation of the cysteine in a processed lipoprotein (12). An 18-residue signal peptide is also consistent with the results of globomycin treatment shown in Fig. 1; the predicted size of the peptide is approximately the difference in sizes between the two bands in the globomycin-treated cells.

Comparative sequence analysis of *vmp33*. Primer extension analysis (Fig. 4) identified the probable transcriptional site as a T, which was designated position +1 (Fig. 3). At positions 13 to −7 was the likely −10 promoter element (TATTAAT). A short inverted repeat of this sequence included the transcriptional start site. The tetranucleotide TTGA typical of the −35 element of a σ^{70}-type promoter began at position −35. Ending 7 nucleotides upstream of the probable −35 element was a sequence in which 16 of 19 nucleotides were T; a highly A-T-rich region extended back to position −76. At positions +7 to +11 was the consensus ribosomal binding site GGAGG (36). A possible rho-independent terminator was identified between positions 739 and 763.

The deduced amino acid sequence of Vmp33 was compared with sequences of Osp proteins of the Lyme disease borrelias (11, 16, 27, 39) and the other Vmp proteins of *B. hermsii* (14, 31). The PILEUP algorithm was used to create the dendrogram shown in Fig. 5. The length of the branches of the tree indicate the similarity score (18), the scale for which is provided above the tree. The analysis revealed that Vmp33, Vmp3, and the two OspC proteins were more closely related to

```
      -100                  -80                  -60
ATTTGTATAGATATTGATAATATTTTTGTTTTGTAAAATTATTTACGAATTTTTTTATAA
       -40                  -20                  +1
TTTTTTTTAAAAACTTGAAAAGTAAAATTTTTGCACTATTAATGTATGTTAATAAGGAGG
                                          ----->      <-----  RBS
       20                   40                   60
CACATAATAATGAAGAAGAATACATTAAGTGCGATATTGATGACTTTATTTTTATTTATA
         M  K  K  N  T  L  S  A  I  L  M  T  L  F  L  F  I
       80                  100                  120
TCTTGTAATAATGGAGGCCCAGAGCTTAAAAAGGGTGAAGTAACTAAATCTGATGGAACA
S  C  N  N  G  G  P  E  L  K  K  G  E  V  T  K  S  D  G  T
      140                  160                  180
GTTCTTGATTTATCAAAAATAAGTGCAAATATAAAAAACGCTGTTACTTTTGCTGCAAGT
V  L  D  L  S  K  I  S  A  N  I  K  N  A  V  T  F  A  A  S
      200                  220                  240
GTTCAAGAAGTAGAGACCTTAGTTAAGTCTATTGATGAGCTTGCTAAAGCTATTGGACAG
V  Q  E  V  E  T  L  V  K  S  I  D  E  L  A  K  A  I  G  Q
      260                  280                  300
AAAGTTAATGCAGATGGTCTTACTGCTGAAGCAAATAAAAAATGATTCATTAGTTGCAGGG
K  V  N  A  D  G  L  T  A  E  A  N  K  N  D  S  L  V  A  G
      320                  340                  360
GTATATCAATTAATATCAGATGTACAGGGTAAGTTGACAAAATTGGAAATTGGGGCCAGC
V  Y  Q  L  I  S  D  V  Q  G  K  L  T  K  L  E  I  G  A  S
      380                  400                  420
AAGTTTGCTGGACTAAAGGAAAAGGTTGTTGCTGCTAAGAAGGGAAGTGATGATTTCTTA
K  F  A  G  L  K  E  K  V  V  A  A  K  K  G  S  D  D  F  L
      440                  460                  480
ACCAAAGTAAAAGCGCAGCATAATAATCTTGGTCAAAGTGCTGAAGCCCCAAAGGCTATA
T  K  V  K  A  Q  H  N  N  L  G  Q  S  A  E  A  P  K  A  I
      500                  520                  540
AAGAAAGGTAATGCTGATAGTACTAAGGGAGCTGAAGAACTTGGTAAGTTAAATACAGCA
K  K  G  N  A  D  S  T  K  G  A  E  E  L  G  K  L  N  T  A
      560                  580                  600
ATAGATGAGTTGTTAACCGCTGCTAAAGATGCAGTAGAAGCTGCAATTGCGGACGTTACA
I  D  E  L  L  T  A  A  K  D  A  V  E  A  A  I  A  D  V  T
      620                  640                  660
GCTGCTCCTGCTAAGCCAGCGACCCCTGTTAAACCTTAACTAAGGTATATATAATTTTAA
A  A  P  A  K  P  A  T  P  V  K  P  *              *     *
      680                  700                  720
ATATTATTATAAGATTAAATTTTAATCAAAATGTAAGTAACTGGAAAAATAAAGTCAATA
      740                  760                  780
AATAGAAAGCTAAGAGCACTATGCTCTTAGCTTCTGATATTATTTAAGACCATCTTTTAT
       >>>>>>>>>>>   <<<<<<<<<<<
      800                  820                  840
CTTTACTTTATTTACTTTCCCAGCGTTTAGATTAAGATTATATTTTTTATTGTATAAGTT
```

FIG. 3. DNA sequence and deduced amino acid sequence of the expressed *vmp33* gene of *B. hermsii* (HS1 strain). Nucleotides are numbered relative to the transcription initiation site (+1). Significant features of the DNA sequence (in order from the 5′ end) include the following: short inverted repeats (- - →← - -) around the transcriptional start site, a consensus ribosome binding sequence (RBS), three termination codons at the 3′ end of the open reading frame, and complementary regions consistent with hairpin loop formation (>>>> and <<<<). The open reading frame, which begins at position +21, encodes a 211-amino-acid polypeptide with a hydrophobic leader sequence and a consensus signal peptidase II cleavage site (amino acids 1 to 19). The underlined amino acids correspond precisely to the sequences obtained from proteolytic cleavage fragments of the Vmp33 protein.

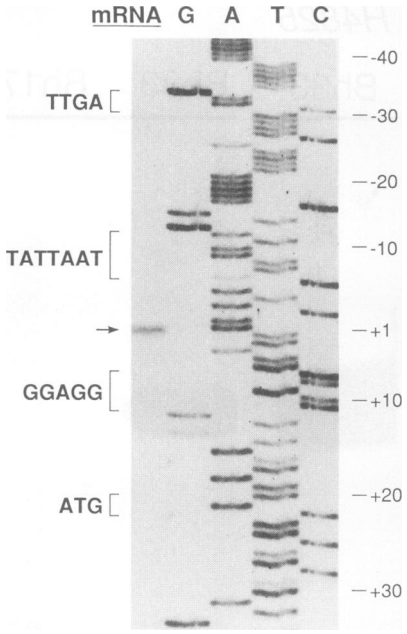

FIG. 4. Primer extension analysis of the 5′ terminus of the expressed *vmp33* gene of *B. hermsii* HS1. The oligonucleotide used as primer was 5′-TTC TTG AAC ACT TGC AGC-3′, which is the reverse complement of nucleotides 183 to 200 in Fig. 2. The primer was annealed to the mRNA, and the DNA was extended with reverse transcriptase. The resulting primer extension product is in the leftmost lane and is indicated by an arrow. A dideoxynucleotide sequencing reaction using the same primer and the vmp33 gene as template was carried out in parallel to the primer extension analysis, yielding a sequence complementary to the coding strand. The locations of the upstream −35 (TTGA) and −10 (TATTAAT) promoter elements, a consensus ribosomal binding sequence (GGAGG), and the putative start codon (ATG) are indicated. For convenience, these designations refer to sequences that would be on the strand opposite to that shown under lanes G, A, T, and C.

each other than they were to other outer membrane proteins of the same species. In other words, Vmp33 had greater sequence identity with OspC of *B. burgdorferi* than it did with Vmp7 or Vmp17 of the same strain. Likewise, OspC was more similar to Vmp3 and Vmp33 of *B. hermsii* than it was to OspA, -B, or -D of the same strain.

The aligned DNA sequence identities determined by the GAP algorithm between the pairs *vmp33-ospC*, *vmp33-vmp3*, and *vmp3-ospC* were 61, 70, and 62%, respectively. The corresponding amino acid sequence identities were 48, 55, and 46%. Calculated sizes of Vmp33 and OspC of strain B31 cleaved of their presumed signal peptides are 19,620 and 20,287, respectively. The predicted isoelectric points of the processed proteins are 8.4 for Vmp33 and 7.4 for OspC of strain B31. A comparison of the mole percent compositions of the processed proteins of Vmp33 and OspC of strain B31 is given in the Table. Amino acid compositions of the two proteins are similar. Of note are the high lysine content and the absence of arginine and tryptophan in both proteins.

When the sequences of Vmp33, Vmp3, and the OspC of strain B31 were examined more closely, a distinction between Vmp33 and other known Vmp proteins was found (Fig. 6). The sequences of Vmp3, -7, -17, -21, -24, and -25, while different for most of their lengths, have the same N termini for the unprocessed protein, MRKRISAII (14, 31). In contrast, the deduced N-terminal sequence of the unprocessed Vmp33 was identical for 21 residues to the OspC of *B. burgdorferi* and other Lyme disease borrelias (39). When this 21-residue conserved sequence was used as the search sequence with a BLAST survey with a percent accepted mutations (PAM) setting of 300 of the Swiss Protein and PIR protein data bases, the closest match (Poisson probability, 0.08) was with the 11 N-terminal residues (MKKKIISAILM) of the precursor of the secreted 45-kDa protein of *Lactococcus lactis* (37).

At its C-terminal end, from position 164 to 211, Vmp33 was less similar to OspC than it was to Vmp3 but more similar to OspC than was Vmp3 (Fig. 6). For this region, the numbers of identical residues between the OspC-Vmp33, Vmp33-Vmp3, and OspC-Vmp3 pairs were 19, 26, and 10, respectively. Moreover, the ends of Vmp33 and OspC, like their beginnings, were the same. The dipeptide lysine-proline also terminates the OspC proteins of other Lyme disease borrelias (39).

Plots of the predicted hydrophilic (negative values) and hydrophobic (positive values) regions of OspC of strain B31, Vmp33, and Vmp3 are shown in Fig. 7. The patterns are

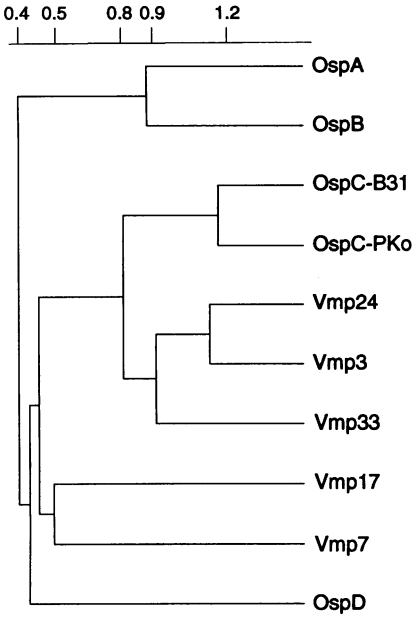

FIG. 5. Dendrogram of clustering sequence relationships of Vmp and Osp proteins of *B. hermsii* and *B. burgdorferi*. See the text for sources of sequences. Amino acid sequences of unprocessed Osp and Vmp proteins were compared by the PILEUP algorithm, and the order of the pairwise alignments was plotted (15, 18). A scale of the similarity scores for the branches is shown above the plot.

similar. Notable are the hydrophobic N termini, the abrupt transition over residues 90 to 100 from a highly hydrophilic region to a hydrophobic region, and hydrophilic C termini in the three proteins.

DISCUSSION

Relapsing fever and Lyme disease borrelias have surface-exposed lipoproteins of 20 kDa that are members of the same protein family. The evidence for this was antigenic cross-reactivity between Vmp33 and OspC, primary sequence identities of 40 to 50%, and similar predicted patterns of hydrophilicity and hydrophobicity. We have also found in partial amino acid sequences that the species *Borrelia turicatae* has proteins of this family (14a).

TABLE 1. Deduced amino acid compositions of OspC of *B. burgdorferi* B31 and Vmp33 of *B. hermsii* HS1

Residue	Mol %	
	OspC-B31[a]	Vmp33
Ala	10.9	16.2
Cys	1.0	0.5
Asp	5.2	6.3
Glu	9.9	6.8
Phe	1.0	1.6
Gly	7.3	7.9
His	1.6	0.5
Ile	5.2	4.7
Lys	15.1	13.6
Leu	10.4	8.9
Met	0.5	0.0
Asn	7.8	5.8
Pro	2.1	3.1
Gln	1.0	3.1
Arg	0.0	0.0
Ser	8.3	5.8
Thr	6.8	6.3
Val	5.2	8.4
Trp	0.0	0.0
Tyr	0.5	0.5

[a] Data for OspC of strain B31 were obtained from reference 40.

```
          1                                                  49
OspC-B31  MKKNTLSAILMTLFL.FISCNNSGKDGNTSANSA.DESVKGPNLTEISKKI
          |||||||||||||||| |||||| |          | |        ||‡‡|
Vmp33     MKKNTLSAILMTLFL.FISCNNGGPELKKGEVTKSDGTVLDLSK..ISANI
          | |   ||| ||||  | ||||||||||  || |||||||||  |  ||‡‡|
Vmp3      MRKRI.SAIIMTLFMVFMSCNNGGPELKSDEVAKSDGTVLDLAK..ISKKI

          164                                                211
OspC-B31  GTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT.SPVVAESPKKP*
          ‡|| ‡||||||||       | ‡|||       ‡‡‡| ‡|‡ ‡‡ | ||
Vmp33     STK..GAEELGKLNTAIDELLTAAKDAVEAAIADVTAAPAKPATPVKP*
          ‡|| ‡|| || ||||||||||‡|| ||||  |‡‡‡||‡ ‡ ‡‡ |
Vmp3      GTKDKGAAELIKLNTAIDELLKAANDAVETVIKELTAS.VKAEKPSQNN*
```

FIG. 6. Alignment of N-terminal and C-terminal amino acids of unprocessed OspC protein of *B. burgdorferi* B31 and Vmp33 and Vmp3 proteins of *B. hermsii* HS1. Single-residue gaps (.) were introduced for optimal alignment. Identity in amino acid residues at a given position between adjacent rows, i.e., OspC-B31-Vmp33 or Vmp33-Vmp3, is indicated by |. Identity in amino acid residues at a given position between OspC of strain B31 and Vmp3 is indicated by ‡.

The overall DNA identity between *B. hermsii* and *B. burgdorferi* has been estimated by hybridization to be 30 to 44%, a divergence that is consistent with their designations as different species (reviewed in reference 6). The amount of actual DNA sequence identity between genes for the periplasmic protein flagellin is 87% for *B. burgdorferi* and *B. hermsii* (17, 26, 29). As would be expected of a conserved, structurally important protein like flagellin, the amino acid identity between the flagellins of these two species is higher at 92% than the nucleotide identity. In contrast, between Vmp33 and OspC and their genes, there was greater nucleotide identity (61%) than amino acid sequence identity (48%). This lower sequence similarity at the amino acid level than at the nucleotide level is an indication that the 20-kDa proteins are immune targets in their respective cells. Under the selective pressure from the immune response of the vertebrate hosts, these proteins of *B. hermsii* vary in their antigenicities even within clonal populations (7, 9). In *B. hermsii*, the small Vmp's, like the large ones, are determinants of serotype specificity (31). The OspC proteins of Lyme disease borrelias also vary, by as much as 20 to 25%, in amino acid sequence (39). However, there is usually another major protein, such as OspA, expressed at the same time, and consequently, the contribution of this family of proteins in determining serotype identity among Lyme disease borrelias is less clear. Nevertheless, there is evidence that development of an antibody response against OspC protects animals against experimental challenge infection (30), and in this sense OspC is an immunodominant protein.

Another similarity between OspC and Vmp33 is their heightened production when expression of other abundant surface proteins is decreased or absent. In the B31 strain of *B. burgdorferi*, expression of OspC was greatest when the plasmid bearing the genes for OspA and OspB was lost from the cells (34). Vmp33, like the other Vmp proteins of *B. hermsii*, is the single abundant outer membrane protein of cells and is always expressed. In another respect, though, serotype 33 differs from other serotypes of *B. hermsii*. Vmp33 has been identified only in cultures that have been passaged in broth several times in the laboratory (9). Although serotype 33 remains infectious for mice, there is a much lower frequency of relapse of spirochet-

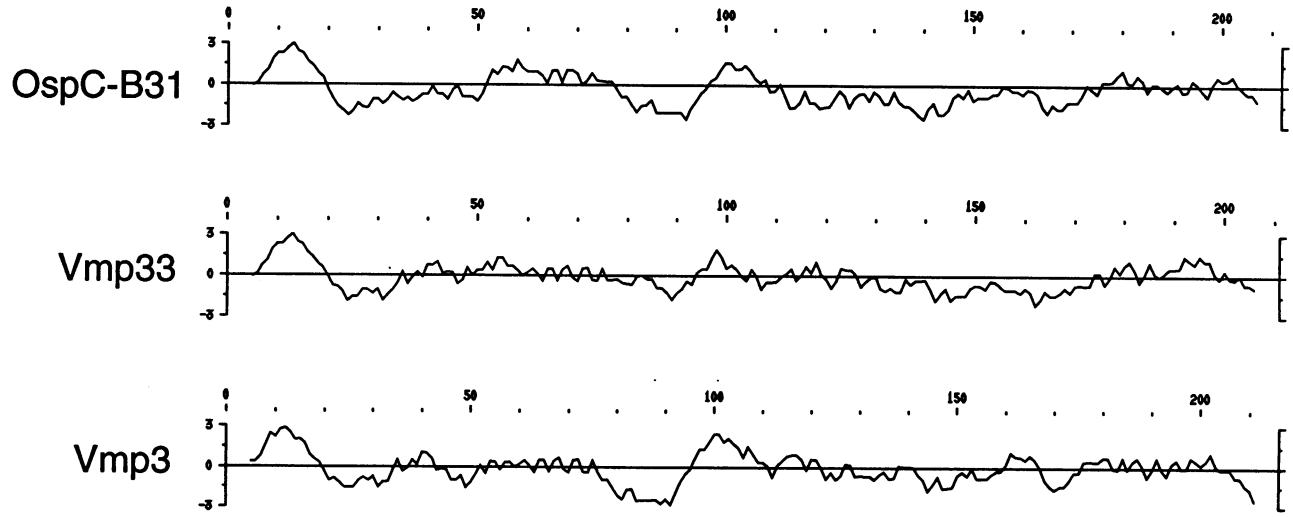

FIG. 7. Hydrophobicity (0 to +3) and hydrophilicity (−3 to 0) profiles by PEPPLOT Kyte-Doolittle algorithm (18) of OspC of *B. burgdorferi* B31 and Vmp33 and Vmp3 of *B. hermsii* HS1. The scale at the top of each plot indicates the amino acid position.

emia in mice infected with serotype 33 than has been observed with other serotypes (3a, 7). The finding that unprocessed Vmp33 differs from other Vmp proteins at its N terminus suggests that Vmp33 in *B. hermsii*, like OspC in *B. burgdorferi* (34), is expressed from a locus other than that expressed from other major lipoproteins characterized to date (5, 21, 31). Consistent with this conclusion was the finding of a promoter region for *vmp33* that was different from the expression site previously identified for *vmp3*, *-7*, *-17*, *-21*, *-24*, and *-25* (5, 31). However, if the expression site for *vmp33* is different from that for other *vmp* genes, it is not clear why another Vmp protein is not expressed simultaneously. In contrast to the missing *ospAB* operon in some $OspA^-$ $OspB^-$ $OspC^+$ mutants of *B. burgdorferi* (32, 34), the genes for other Vmp proteins are not lost from serotype 33 cells (25). Other mutants and strains of Lyme disease borrelias express OspC along with OspA (34, 39). The findings suggest that, in *B. hermsii*, expression at one locus appears to influence the state of activation at another potential expression site. It appears that, when the previously described, and apparently dominant, expression site of serotypes 3, 7, 17, 21, 24, and 25 is active (14, 21, 31), expression of the *vmp33* gene is repressed at the level of transcription or otherwise altered posttranscriptionally.

ACKNOWLEDGMENTS

We thank Tom Schwan of the Rocky Mountain Laboratories and Bettina Wilske of the Pettenkofer Institute, University of Munich, for providing reagents and helpful discussions and Cindy Freitag for her early contributions to this project.

This work was supported by grants AI24424 and AI29731 from the National Institutes of Allergy and Infectious Diseases, grant AR41507 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases, and grant 07922 from the Swedish Medical Research Council. We thank Per-Ingvar Ohlsson, University of Umeå, Sweden, for performing amino acid sequence analysis under support of grant 07XAC from the Swedish Cancer Society.

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**:403–410.
2. **Barbour, A. G.** 1984. Isolation and cultivation of Lyme disease spirochetes. Yale J. Biol. Med. **78**:521–525.
3. **Barbour, A. G.** 1985. Clonal polymorphism of surface antigens in a relapsing fever *Borrelia* species, p. 235–245. *In* G. G. Jackson (ed.), Bayer Symposium VII: The pathogenesis of bacterial infections. Springer-Verlag, Heidelberg.

3a. **Barbour, A. G.** Unpublished data.

4. **Barbour, A. G., O. Barrera, and R. C. Judd.** 1983. Structural analysis of the variable major proteins of *Borrelia hermsii*. J. Exp. Med. **158**:2127–2140.
5. **Barbour, A. G., N. Burman, C. J. Carter, T. Kitten, and S. Bergström.** 1991. Variable antigen genes of the relapsing fever agent *Borrelia hermsii* are activated by promoter addition. Mol. Microbiol. **5**:489–493.
6. **Barbour, A. G., and S. F. Hayes.** 1986. Biology of *Borrelia* species. Microbiol. Rev. **50**:381–400.
7. **Barbour, A. G., and H. G. Stoenner.** 1985. Antigenic variation of *Borrelia hermsii*, p. 123–135. *In* I. Herskowitz and M. I. Simon (ed.), Genome rearrangement. Alan R. Liss, Inc., New York.
8. **Barbour, A. G., S. L. Tessier, and S. F. Hayes.** 1984. Variation in a major surface protein of Lyme disease spirochetes. Infect. Immun. **45**:94–100.
9. **Barbour, A. G., S. L. Tessier, and H. G. Stoenner.** 1982. Variable major proteins of *Borrelia hermsii*. J. Exp. Med. **156**:1312–1324.
10. **Barstad, P. A., J. E. Coligan, M. G. Raum, and A. G. Barbour.** 1985. Variable major proteins of *Borrelia hermsii*. Epitope mapping and partial sequence analysis of CNBr peptides. J. Exp. Med. **161**:1302–1314.
11. **Bergström, S., V. G. Bundoc, and A. G. Barbour.** 1989. Molecular analysis of linear plasmid-encoded major surface proteins, OspA and OspB, of the Lyme disease spirochaete *Borrelia burgdorferi*. Mol. Microbiol. **3**:479–486.
12. **Brandt, M. E., B. S. Riley, J. D. Radolf, and M. V. Norgard.** 1990. Immunogenic integral membrane proteins of *Borrelia burgdorferi* are lipoproteins. Infect. Immun. **58**:983–991.
13. **Bundoc, V. G., and A. G. Barbour.** 1989. Clonal polymorphisms of outer membrane protein OspB of *Borrelia burgdorferi*. Infect. Immun. **57**:2733–2741.
14. **Burman, N., S. Bergström, B. I. Restrepo, and A. G. Barbour.** 1990. The variable antigens Vmp7 and Vmp21 of the relapsing fever bacterium *Borrelia hermsii* are structurally analogous to the VSG proteins of the African trypanosome. Mol. Microbiol. **4**:1715–1726.

14a. **Cadavid, D., and A. G. Barbour.** Unpublished data.

15. **Feng, D.-F., and R. F. Doolittle.** 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J. Mol. Evol. **35**:351–360.
16. **Fuchs, R., S. Jauris, F. Lottspeich, V. Preac-Mursic, B. Wilske, and E. Soutschek.** 1992. Molecular analysis and expression of a

Borrelia burgdorferi gene encoding a 22 kDa protein (pC) in Escherichia coli. Mol. Microbiol. **6**:503–509.
17. **Gassmann, G. S., M. Kramer, U. B. Göbel, and R. Wallich.** 1989. Nucleotide sequence of a gene encoding the Borrelia burgdorferi flagellin. Nucleic Acids Res. **17**:3590.
18. **Genetics Computer Group.** 1991. Program manual for the GCG package, version 7. Genetics Computer Group, Madison, Wis.
19. **Hinnebusch, J., and A. G. Barbour.** 1992. Linear- and circular-plasmid copy numbers in *Borrelia burgdorferi*. J. Bacteriol. **174**: 5251–5257.
20. **Jonsson, M., L. Noppa, A. G. Barbour, and S. Bergström.** 1992. Heterogeneity of outer membrane proteins in *Borrelia burgdorferi*: comparison of *osp* operons of three isolates of different geographic origins. Infect. Immun. **60**:1845–1853.
21. **Kitten, T., and A. G. Barbour.** 1989. Juxtaposition of expressed variable antigen genes with a conserved telomere in the bacterium *Borrelia hermsii*. Proc. Natl. Acad. Sci. USA **87**:6077–6081.
22. **Kitten, T., and A. G. Barbour.** 1992. The relapsing fever agent *Borrelia hermsii* has multiple copies of its chromosome and linear plasmids. Genetics **132**:311–324.
23. **Marconi, R. T., and C. F. Garon.** 1992. Phylogenetic analysis of the genus *Borrelia*: a comparison of North American and European isolates of *Borrelia burgdorferi*. J. Bacteriol. **174**:241–244.
24. **Matsudaira, P.** 1987. Sequence from picomole quantities of protein electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262**:10035–10038.
25. **Meier, J. T., M. I. Simon, and A. G. Barbour.** 1985. Antigenic variation is associated with DNA rearrangements in a relapsing fever spirochete. Cell **41**:403–409.
26. **Noppa, L., A. Sadziene, A. G. Barbour, and S. Bergström.** 1992. Molecular characterization of the flagellin genes and gene products from *Borrelia anserina*, *B. burgdorferi*, *B. crocidurae*, and *B. hersmii*. abstr. A27, p. 154. Abstr. 5th Int. Conf. Lyme Borreliosis. 1992.
27. **Norris, S. J., C. J. Carter, J. K. Howell, and A. G. Barbour.** 1992. Low-passage-associated proteins of *Borrelia burgdorferi*: characterization and molecular cloning of OspD, a surface-exposed, plasmid-encoded lipoprotein. Infect. Immun. **60**:4662–4672.
28. **Ochman, H., M. M. Medora, D. Garza, and D. L. Hartl.** 1990. Amplification of flanking sequences by inverse PCR, p. 219–227. *In* M. A. Innis, D. H. Gilfand, J. J. Sninsky, and T. J. White (ed.), PCR protocols. Academic Press, Inc., San Diego.
29. **Picken, R. N.** 1992. Polymerase chain reaction primers and probes derived from flagellin sequences for specific detection of the agents of Lyme disease and North American relapsing fever. J. Clin. Microbiol. **30**:99–114.
30. **Preac-Mursic, V., B. Wilske, E. Patsouris, S. Jauris, G. Will, E. Soutschek, S. Reinhardt, G. Lehnert, E. Klockmann, and P. Mehraein.** 1992. Active immunization with pC protein of *Borrelia burgdorferi* protects gerbils against *Borrelia burgdorferi* infection. Infection **20**:342–349.
31. **Restrepo, B. I., T. Kitten, C. J. Carter, D. Infante, and A. G. Barbour.** 1992. Subtelomeric expression regions of *Borrelia hermsii* linear plasmids are highly polymorphic. Mol. Microbiol. **6**:3299–3311.
32. **Šadžiene, A., P. A. Rosa, P. A. Thompson, D. M. Hogan, and A. G. Barbour.** 1992. Antibody-resistant mutants of *Borrelia burgdorferi*: in vitro selection and characterization. J. Exp. Med. **176**:799–809.
33. **Šadžiene, A., P. A. Thompson, and A. G. Barbour.** 1993. In vitro inhibition of *Borrelia burgdorferi* growth by antibodies. J. Infect. Dis. **167**:165–172.
34. **Šadžiene, A., B. Wilske, M. S. Ferdows, and A. G. Barbour.** 1993. The cryptic *ospC* gene of *Borrelia burgdorferi* B31 is located on a circular plasmid. Infect. Immun. **61**:2192–2195.
35. **Schwan, T. G., M. E. Schrumpf, R. H. Karstens, J. R. Clover, J. Wong, M. Daugherty, M. Struthers, and P. A. Rosa.** 1993. Distribution and molecular analysis of Lyme disease spirochetes, *Borrelia burgdorferi*, isolated from ticks throughout California. J. Clin. Microbiol. **31**:3096–3108.
36. **Stormo, G. D., T. D. Schneider, and L. M. Gold.** 1982. Characterization of translational initiation sites in *E. coli*. Nucleic Acids Res. **10**:2971–2996.
37. **van Asseldonk, M., G. Rutten, M. Oteman, R. J. Siezen, W. M. De Vos, and G. Simons.** 1990. Cloning of usp45, a gene encoding a secreted protein from Lactococcus lactis subsp. lactis MG1363. Gene **95**:155–160.
38. **von Heijne, G.** 1983. Patterns of amino acids near signal sequence cleavage sites. Eur. J. Biochem. **133**:17–21.
39. **Wilske, B., V. Preac-Mursic, S. Jauris, I. Pradel, E. Soutschek, E. Schwab, and G. Wanner.** 1993. Immunological and molecular polymorphism of OspC: an immunodominant major outer surface protein of *Borrelia burgdorferi*. Infect. Immun. **61**:2182–2191.
40. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains. Gene **33**:103–119.