

UNIVERSITY OF WISCONSIN-LA CROSSE

Graduate Studies

OUTER SURFACE PROTEIN (OSP) F SUBFAMILY ANTIBODY RESPONSES  
DURING LYME DISEASE

A Manuscript Style Thesis Submitted in Partial Fulfillment of the Requirements for the  
Degree of Master of Science in Microbiology

Ashley Werner

College of Science and Health  
Clinical Microbiology


December, 2018

OUTER SURFACE PROTEIN (OSP) F SUBFAMILY ANTIBODY RESPONSES  
DURING LYME DISEASE


By Ashley Werner

We recommend acceptance of this thesis in partial fulfillment of the candidate's requirements for the degree of Microbiology-Clinical Microbiology Concentration

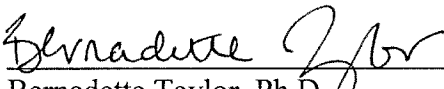
The candidate has completed the oral defense of the thesis.

  
\_\_\_\_\_  
Steven Callister, Ph.D.  
Thesis Committee Chairperson


11/15/2018  
Date

  
\_\_\_\_\_  
Steven Lovrich, Ph.D.  
Thesis Committee Member

11/15/2018  
Date

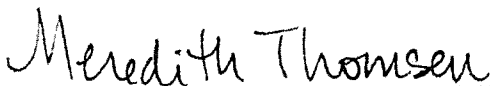
  
\_\_\_\_\_  
Bernadette Taylor, Ph.D.  
Thesis Committee Member

11/16/2018  
Date

  
\_\_\_\_\_  
Peter Wilker, Ph.D.  
Thesis Committee Member

11/16/2018  
Date

Thesis accepted

  
\_\_\_\_\_  
Meredith Thomsen, Ph.D.  
Graduate Studies Director

12-13-2018  
Date

## ABSTRACT

Werner, A.M. Outer surface protein (Osp) F subfamily antibody responses during Lyme disease. MS in Clinical Microbiology, December 2018, 58 pp. (S.M. Callister)

Lyme disease is the most common tickborne illness in the United States. The CDC currently recommends a two-test system comprised of an enzyme linked immunosorbent assay followed by a Western blot for laboratory diagnosis. However, the approach is cumbersome and expensive, so identifying additional proteins that can be used to detect antibody responses associated exclusively with Lyme disease remains an important area of research. In this study, the genes encoding outer surface protein (Osp) F subfamily members OspF, OspEF-related protein G (ErpG), ErpK, and ErpL were cloned, and the resultant recombinant proteins were used to evaluate the antibody responses in well-characterized human Lyme disease sera. Although significant numbers of the Lyme disease sera contained antibodies that bound ErpG, ErpK, and/or ErpL, the responses were without an apparent pattern. In contrast, OspF antibodies were detected in a majority of patients with disseminated infection including 81% of patients with joint pain or swelling and each patient with cardiac abnormalities (3/3) or neuritis/neuropathy (2/2). Therefore, the findings failed to provide support for the diagnostic utility of the Erps, but additional studies to evaluate the utility of OspF for Lyme disease diagnostics are warranted.

## ACKNOWLEDGEMENTS

First, I would like to thank my major advisor, Dr. Steve Callister, for his support, advice with my research, and help with the writing process. I would also like to thank Dr. Steve Lovrich for his assistance with technical training and study design. Additionally, I would like to thank the other members of my committee, Dr. Peter Wilker and Dr. Bernadette Taylor, for their guidance throughout the course of my studies. Lastly, I would like to acknowledge the generous financial contribution of the Gundersen Medical Foundation.

# TABLE OF CONTENTS

	PAGE
LIST OF FIGURES.....	vii
LIST OF TABLES.....	viii
INTRODUCTION.....	1
STUDY RATIONALE AND SPECIFIC OBJECTIVES.....	7
METHODS AND MATERIALS.....	10
Normal sera.....	10
Lyme disease sera.....	10
Chart review.....	10
Western blotting.....	11
Selection, synthesis, and recovery of genes.....	11
Insertion into expression vector and propagation of plasmids.....	14
Transformation of BL21 Star cells.....	16
Purification of proteins.....	16
Sodium dodecyl sulfate-polyacrylamide gel electrophoresis.....	17
Enzyme linked immunosorbent assay.....	17
Statistics.....	19

RESULTS.....	21
Recovery of recombinant proteins.....	21
Characterization of Lyme disease sera.....	29
Detection of ErpG, ErpK, and ErpL antibodies.....	31
Detection of OspF antibodies.....	33
DISCUSSION.....	35
REFERENCES.....	39
APPENDICES.....	45
Truncated sequences of <i>ospF</i> , <i>erpG</i> , <i>erpK</i> , and <i>erpL</i> .....	45
Vector pET100 /D-TOPO sequence data.....	48

## LIST OF FIGURES

FIGURE	PAGE
1. Schematic illustration and genetic sequence of pET100 vector. The insertion site is shaded.....	15
2. Schematic illustration of cloning strategy.....	22
3. Detection of amplified genes by agarose gel electrophoresis. Lane 1 contains 50 bp-10 kbp Perfect DNA Marker, lane 2 contains <i>ospF</i> fragment, lane 3 contains <i>erpG</i> fragment, lane 4 contains <i>erpK</i> fragment, and lane 5 contains <i>erpL</i> fragment.....	23
4. Comparison of recovered <i>ospF</i> DNA sequence to the requested sequence. Vertical dashes indicate exact matches.....	24
5. Comparison of recovered <i>erpG</i> DNA sequence to the requested sequence. Vertical dashes indicate exact matches.....	25
6. Comparison of recovered <i>erpK</i> DNA sequence to the requested sequence. Vertical dashes indicate exact matches.....	26
7. Comparison of recovered <i>erpL</i> DNA sequence to the requested sequence. Vertical dashes indicate exact matches.....	27
8. Detection of recombinant proteins by SDS-PAGE. Lane 1 contains OspF, lane 2 contains ErpG, lane 3 contains ErpK, and lane 4 contains ErpL.....	28

## LIST OF TABLES

TABLE	PAGE
1. Primer sequences and expected amplicon sizes.....	13
2. Significant optical density (OD) cut-off values for ErpK, ErpL, ErpG or OspF ELISA.....	20
3. Primary clinical abnormalities of patients with Lyme disease confirmed by CDC Western blot criteria.....	30
4. Detection of ErpG, ErpK, or ErpL antibodies in Lyme disease sera.....	32
5. Detection of OspF antibodies in Lyme disease sera.....	34

## INTRODUCTION

Lyme disease was first described in 1975 after Dr. Allen Steere noted an unusual epidemic of juvenile rheumatoid arthritis in a group of children residing in Lyme, Connecticut (1). However, the affected children lived in wooded areas, had rashes, and reported tick bites, which raised suspicion of an infectious cause. After further study, a spirochete was recovered from the midguts of captured *Ixodes scapularis* ticks (2), and the sera from the patients contained antibodies that bound the spirochetes (3). The newly-described organism was subsequently named *Borrelia burgdorferi* and, following discovery in the United States (US), similar spirochetes were isolated from both *I. ricinus* ticks and patients in Europe with symptoms that mimicked the illness seen in the US patients.

The initial impression, therefore, was that the organisms affecting the European patients were also *B. burgdorferi*, but later studies showed that, in addition to illness from *B. burgdorferi*, Lyme disease in European and Asian patients could also be caused by transmission of genetically-distinct *B. afzelii* or *B. garinii*. Since then, researchers (4-6) have confirmed that multiple *Borrelia* spp. cause Lyme disease, and the spirochetes are now classified as members of the *B. burgdorferi* sensu lato complex until genetic characterization has been completed. Despite the great genetic diversity, however, the most common causative agent of Lyme disease in the US remains *B. burgdorferi*, which is now recognized officially as *B. burgdorferi* sensu stricto (ss). Since its discovery,

Lyme disease has become the most common tickborne infection in North America, and hyperendemic foci of *B. burgdorferi* ss-infected *I. scapularis* ticks are located throughout the midwestern and northeastern regions of the US.

The illness is categorized by severity of symptoms and duration of infection, and formally-recognized categories include early localized, early disseminated or late disseminated stages. Early localized infection manifests primarily as non-specific constitutional abnormalities such as fatigue, headache, fever, arthralgia, and myalgia. In addition, approximately 80% of patients develop a rash, termed erythema migrans (EM) (7), that is typically an expanding red ring with central clearing. In fact, the EM lesion is associated almost exclusively with early Lyme disease, so most clinicians consider a characteristic EM a pathognomonic sign of the illness (8).

If the early localized infection is not treated appropriately, the organisms disseminate to a variety of other locations. Therefore, the symptoms of early disseminated infection depend on where the spirochetes have colonized, but well-characterized complications include secondary EM lesions (9), facial palsy (10), or cardiac rhythm abnormalities (11). In addition to these possibilities, however, the most common complaint in US Lyme disease patients is likely arthralgia, presumably because *B. burgdorferi* ss organisms have predilection for joint tissue (12). If a disseminated infection remains untreated or is treated inappropriately, symptoms of late stage illness are more severe abnormalities that include heart blockage (13), a dementia-like syndrome (14) or, most commonly, recurrent joint swelling (8).

After the spirochetal microorganisms were identified as the causative agent of Lyme disease and researchers confirmed the organisms sequestered in multiple sites, fears that the infection caused an illness that mimicked syphilis caused by *Treponema pallidum* spirochetes emerged. This was especially worrisome, because the prevalence of syphilis continues to increase worldwide, and morbidity and mortality associated with the illness are significant. In addition, the syphilis spirochetes are difficult to eliminate by antibiotic therapy since the organisms readily cross the blood-brain barrier (15). Moreover, since ticks infected with *Borrelia* spp. are endemic throughout many regions of the US, the risk of acquiring Lyme disease would likely be much more significant.

Not surprisingly, therefore, the initial strategy for laboratory testing emphasized the need to detect the illness early after infection. However, directly confirming the infection by recovering the organisms remains difficult, because the culture medium is difficult to make, and appropriate commercial sources of prepared medium are not readily-available. In addition, directly detecting the *Borrelia* spp. DNA by polymerase chain reaction (PCR) technology was not yet developed and, even after development, the test suffers from the inability to obtain samples that reliably contain the spirochetes. Therefore, providing indirect evidence of infection remains the most viable option, and detecting antibodies formed against the spirochetes is the norm.

Primary among the antibody tests were enzyme linked immunosorbent assays (ELISA), indirect fluorescent antibody (IFA) tests, or Western blotting, but each suffered significantly from a gross lack of specificity (16). For example, antibodies produced in patients with rheumatoid arthritis (17), multiple sclerosis (MS) (17), mononucleosis (18), and multiple other infections or conditions often also bound *Borrelia* spp. proteins, and

the cross-reactivity caused considerable confusion. The result was individuals with neurologic illnesses like MS or amyotrophic lateral sclerosis (ALS) would obtain a positive result from a Lyme disease laboratory test and then use the result to lobby for treatment with antibiotics. The non-specific anti-inflammatory properties of some antibiotics (19) subsequently caused the patients to believe they were on the road to recovery, but the patients inevitably relapsed after the treatment was halted. This caused considerable confusion, especially concerning whether additional antibiotic therapy would be helpful. The confusion ultimately caused the Centers for Disease Control and Prevention (CDC) to mandate focused effort to improve the specificity of Lyme disease tests, and the result was adaptation of a two-test system.

The first test was a sensitive, but nonspecific *B. burgdorferi* ELISA that utilized sonicated *B. burgdorferi* organisms to screen for any antibodies that bound the antigens. A diagnosis of Lyme disease was then unlikely if the ELISA yielded a negative result. Lyme disease remained a possibility if a positive result was obtained, however, so the serum was then re-tested by using a Western blot and specific antibodies were required to confirm a diagnosis of the illness. An IgM Western blot was considered positive if antibodies bound at least two of the *B. burgdorferi* ss 24 kDa, 39 kDa, or the 41 kDa protein (20), and a positive IgG blot required antibodies to bind at least five of the 18 kDa, 21 kDa, 28 kDa, 30 kDa, 39 kDa, 41 kDa, 45 kDa, 58 kDa, 66 kDa, or 93 kDa proteins (21).

The two-test regimen resulted in more accurate confirmation of Lyme disease. However, the requirement is cumbersome and expensive, which caused many laboratories to resist implementation. Therefore, researchers have continued to evaluate

alternate diagnostic laboratory testing procedures. Most prominent among these are peptide ELISAs that offer higher specificity by detecting antibodies that bind specific regions of individual *B. burgdorferi* proteins. The most useful have been pepC10 or C6 ELISAs that detect antibodies that bind specific epitopes within OspC (22) or the variable surface expressed (VlsE) protein (23), respectively.

Specifically, the pepC10 ELISA detects antibodies that bind the carboxy terminus of the OspC protein and, since OspC is reliably expressed shortly after the spirochetes invade the mammalian host (24, 25), pepC10 antibodies are produced shortly after infection. A drawback, however, is lack of sensitivity during the earliest stage of Lyme disease because it typically takes several weeks of infection before antibodies are produced. For example, Mathiesen et al. (22) showed that pepC10 antibodies were produced by only approximately 40% of patients with EM. In addition, the antibody response does not persist after the spirochetes disseminate because OspC expression is downregulated as the spirochetes leave the skin (26), and the antigen appears to be processed independent of T-cells (27), which makes the response predominately IgM antibodies (28).

The C6 ELISA complements the pepC10 ELISA, because the test detects antibodies that bind a 26-amino acid peptide based on a conserved region of the VlsE protein, and the VlsE protein is expressed as the spirochetes disseminate from the skin to other tissues (23). This makes the test insensitive during early Lyme disease (29), but highly sensitive after the spirochetes have disseminated. For example, Liang et al. (30) showed that C6 antibodies were produced by 128 (93%) of 138 patients with clinical characteristics

consistent with disseminated Lyme disease. Additionally, although initially marketed separately, Zeus Scientific has recently combined the two assays into a single ELISA.

The combination of the pepC10 and C6 ELISAs has therefore greatly improved the efficiency of laboratory diagnostics, and the new tests offer the added benefit of objective evaluation versus the subjective analysis necessary when interpreting a Western blot. Despite the development of these ELISAs however, the CDC has not to date deemed the improvements sufficient to eliminate their recommendation for the two-test system. Therefore, studies to identify additional Lyme disease-specific antibody responses that may also be useful for diagnostic purposes remain important.

## STUDY RATIONALE AND SPECIFIC OBJECTIVES

The identification of one or more additional proteins or individual epitopes within proteins that reliably induce highly specific antibodies during human Lyme disease remains an important research goal, especially if the antibody response complements the responses detected by using the pepC10 and C6 ELISAs. The OspEF- related (Erp) proteins (31) may contain good candidate antigens because many are expressed on the surface of the spirochetes during human illness (32). However, identifying appropriate candidates that warrant investigation within this family of proteins is complicated by the large number of possibilities. In fact, the family of Erp proteins is so large and diverse that investigators recently reclassified the proteins into subfamilies (33, 34) based on whether the proteins were similar to OspE (OspE-related), OspF (OspF-related), or the proteins have leader peptides that are a combination of both the OspE and OspF leader peptides (OspE/F-like leader peptides or Elps) (35, 36).

Despite the diversity, however, careful review of the scientific literature revealed some study results that suggested several proteins that warrant study for their diagnostic potential. For example, recent studies in mice suggested that antibodies specific for ErpG, ErpK, and ErpL could be used to reliably confirm infection with Lyme disease spirochetes. In addition, the antibody response may yield insight into the location of the infection, because the antibodies specific for each of these proteins appears to differ based on the duration and location of infection. In support, ErpL antibodies are detected

during the first few weeks of infection in mice and ErpK antibodies are not produced until after the spirochetes have disseminated (37, 38). Additionally, the temporal response may be due to predilection for certain tissues as other researchers (35, 39) have also showed that ErpG binds glial cells with high affinity and ErpK specifically binds respiratory epithelial cells. Moreover, the proteins apparently target heparan sulfate (40), a glycosaminoglycan (GAG) (35) that forms proteoglycans by linking with proteins on cell surfaces and also in the extracellular matrix. The structure or composition of the heparan sulfate may also be responsible for the tissue preference.

Another prominent candidate diagnostic antigen is OspF. In support, Magnarelli et al. (41) showed that 5 (25%) of 20 immune sera from patients with EM lesions contained OspF antibodies. More significantly, approximately 90% of patients with early Lyme disease were seropositive when the significant OspF antibody responses were combined with the responses detected by using the pepC10 and C6 ELISAs. In addition, Akin et al. (42) detected OspF antibodies in 80% of immune sera from patients with Lyme arthritis, which suggests the antibody response expands as the spirochetes disseminate. Despite these findings, however, other researchers reported less promising results. For example, Hefty et al. (43) detected OspF antibodies in only 39% of sera from patients with disseminated Lyme disease. In addition, significant sequence heterogeneity among *ospF* within *B. burgdorferi* ss isolates caused others (44, 45) to speculate that the genetic differences would be sufficient to limit its usefulness for diagnostics. The result of these contradictions was that few studies to evaluate the utility of OspF for diagnosing human Lyme disease have since been performed. However, Wagner et al. (46) recently showed that OspF antibodies were a reliable marker of Lyme disease in canines infected

throughout the US, which highlights the necessity for additional investigation into the utility of OspF as a diagnostic antigen for confirming human Lyme disease.

This study was therefore designed to characterize the ErpG, ErpK, ErpL and OspF antibody responses during human Lyme disease. This goal was accomplished by: 1) cloning *erpG*, *erpK*, *erpL*, and *ospF* into an expression vector and preparing ELISAs with each recombinant protein, 2) identifying an appropriate cohort of well-defined human Lyme disease serum samples, 3) reviewing the medical charts of the Lyme disease serum donors to determine the primary clinical abnormalities so that the findings could be used to predict the likely location of infection, and 4) characterizing the IgM and IgG antibody responses specific for each recombinant protein in the Lyme disease sera.

## METHODS AND MATERIALS

**Normal sera.** Normal sera (n = 15) were obtained from healthy adult ( $\geq 18$  years of age) volunteers who resided in the region including and immediately surrounding La Crosse, WI. Each volunteer subject had no previous suspicion (chart review) of Lyme disease. The serum samples were each de-identified, assigned a unique number, and then frozen in 100  $\mu$ l aliquots at  $-20^{\circ}\text{C}$  until used.

**Lyme disease sera.** Archived serum samples from Gundersen Health system patients that tested positive for *B. burgdorferi* VlsE antibodies by a commercial chemiluminescent immunoassay (LIAISON, DiaSorin, Stillwater, MN) were targeted. Some had also already been tested by Western blotting and those that tested positive (n=61) by the IgM or IgG CDC criteria (47) were obtained. In addition, sera with high OD values ( $>3.5$ ) from the chemiluminescent assay that had not been tested previously by Western blotting were tested and those that also contained sufficient levels of antibodies to satisfy the CDC requirement for Western blot positivity (n=50) were also included (47).

**Chart review.** The medical records of serum sample donors were accessed after approval by the Gundersen Health System institutional review board (IRB). Only the primary clinical abnormalities and previous laboratory test results associated with the Lyme disease diagnosis were retrieved.

**Western blotting.** Western blots were performed by using a commercially-available kit (Marblot® IgM and IgG Western Blot kit, Trinity Biotech, Jamestown, NY) according to the manufacturer's instructions. A standardized procedure was used to detect antibodies that bound proteins within the provided nitrocellulose strips, and the band reactivities were scored by comparison to a provided serum band locator strip. A weakly reactive positive control was also included to prevent overdevelopment of the strips.

**Selection, synthesis, and recovery of genes.** Appropriate *ospF*, *erpG*, *erpK*, and *erpL* gene sequences were targeted by analyzing sequences available of the entire gene with a BLAST search. Requirements for selection included thorough characterization and recovery from *B. burgdorferi* 297. The latter requirement was important because the 297 isolate is a well-documented human pathogen (48) with a genetic sequence that closely matches Lyme disease spirochetes recovered in the La Crosse, WI region. Since the promoter sequences of the OspF-subfamily genes are highly similar, the individual gene sequences were chemically synthesized and incorporated into a pUC57 plasmid by a commercial company (Genscript, Piscataway, NJ).

The synthesized genes were then recovered from the pUC57 plasmid using a commercial kit (Phusion High Fidelity PCR Kit, New England Biolabs, Ipswich, MA) and inserted into the expression vector pET100. Briefly, a one  $\mu$ l amount of DNA from a solution of four  $\mu$ g of pUC57 resuspended in 50  $\mu$ l of AE buffer was combined with 20.5  $\mu$ l of master mix that consisted of 13.5  $\mu$ l of PCR grade water, four  $\mu$ l of HF buffer, 0.5  $\mu$ l of 10 mM dNTPs, one  $\mu$ l of forward primer (100  $\mu$ M), one  $\mu$ l of reverse primer (100  $\mu$ M), and 0.5  $\mu$ l of DNA polymerase. To aid in the proper insertion into the pET100 cloning vector, each forward primer had a 5' CACC overhang and the reverse primers

had an additional 3' TAA stop codon (Table 1). Additionally, the primers were designed to create a truncated version of the whole protein (35), which aided in the efficiency of protein recovery by eliminating the N-terminal signal sequence (49). After mixing the suspension, the DNA was amplified by using an initial hold at 98°C for 5 minutes, 40 cycles of denaturation at 98°C for 15 seconds, annealing at 50°C for 30 seconds, elongation at 72°C for 45 seconds, and a final extension at 72°C for 7 minutes (Model 2720, Applied Biosystems, Foster City, CA).

The amplified DNA was separated by electrophoresis (Biorad) in a 2% 1X Tris-acetate-EDTA (TAE) agarose gel. Briefly, the poured gel was allowed to solidify by incubation at room temperature for approximately 30 minutes, 20 µl of amplified DNA combined with four µl of 6X gel loading buffer (Novagen, Temecula, CA) was loaded onto the gel, and the gel was electrophoresed at 100 V for 65 minutes. After electrophoresing, the gel was immersed in 100 ml of distilled water containing 0.7 µg/ml ethidium bromide for 15 minutes at room temperature with gentle rocking, and then evaluated by UV illumination. The size of the recovered DNA was approximated by comparison to perfect DNA marker (Novagen). The portions of the gel that contained the gene of interest were then excised, and the DNA was recovered from the excised portion of the gel by using a commercial kit (MinElute Gel Extraction Kit, Qiagen, Germantown, MD) that bound the DNA to a silica membrane followed by elution in low-salt buffer.

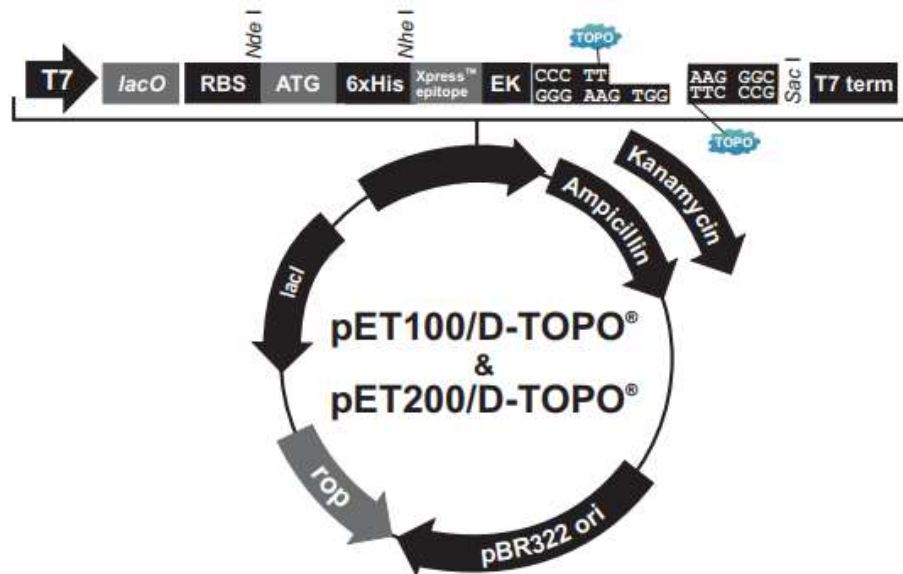
**TABLE 1** Primer sequences and expected amplicon sizes.

Targeted gene	Primer Sequence	Expected Amplicon Size (bp)
<i>ospF</i>	<sup>a</sup> F: CACCGATTTAGAAGGGGCA R: TTATTATTCTTTTTTGA CTCTCCATT	616
<i>erpG</i>	F: CACCGAAGATTTAAAACAAAATGTAA R: TTATTATTTTTTATCTTCTATATTTTGAGGCTC	520
<i>erpK</i>	F: CACCGTAAAAAAAAAGTTTAGAACAAGA R: TTACTATTCTTTTTTATTAGAATCTTTAGATTCTT	688
<i>erpL</i>	F: CACCCTAAAAAATTCAGAACAAAATC R: TTATTATTCTTTTTTATCTTCTTCTATTCCCT	613

<sup>a</sup> F= Forward primer; R= Reverse primer

**Insertion into expression vector and propagation of plasmids.** The purified DNA was then inserted into a pET100 (Invitrogen, Carlsbad, CA) expression vector (Fig. 1). The DNA was initially diluted in elution buffer to a concentration of up to 5 ng/ $\mu$ l and a one  $\mu$ l volume was then combined with one  $\mu$ l of the supplied salt solution and 0.5  $\mu$ l of pET100 vector. During a room temperature incubation, the DNA fragments were incorporated into the linearized plasmid by benefit of an attached topoisomerase.

The pET100 vector containing each DNA insert was transformed into TOP10 *E. coli* cells. Briefly, the TOP10 *E. coli* cells were thawed, pET100 vector was added to the cells, and the mixture was incubated on ice for 20 minutes. After incubation, cells were heat-shocked for 30 seconds at 42°C and immediately transferred back to the ice. The cells were then combined with 250  $\mu$ l of Super Optimal Catabolite (SOC) medium (Invitrogen), followed by a one hour incubation at 37°C. Lastly, a 100  $\mu$ l volume of the SOC medium containing the transformed TOP10 cells was spread onto 2x-YT plates containing 100  $\mu$ g/ml of ampicillin, and the plate was incubated overnight at 37°C. Colonies of transformed TOP10 cells were then picked and cultured in 5ml of 2x-YT containing 100  $\mu$ g/ml of ampicillin overnight at 37°C with shaking. The plasmids were finally recovered by using a commercially available kit (Qiagen) per the manufacturer's instructions, forwarded for commercial sequencing using T7 promoter sequencing primers (Laragen, Culver City, CA), and evaluated by BLAST to confirm the integrity.



```

121 ATAGGCGCCA GCAACCGCAC CTGTGGCGCC GGTGATGCCG GCCACGATGC GTCCGGCGTA GAGGATCGAG ATCTCGATCC
      T7 promoter/priming site
      T7 promoter lac operator
201 CGCGAAATTA ATACGACTCA CTATAGGGGA ATTGTGAGCG GATAACAATT CCCCTCTAGA AATAATTTTG TTAACTTTA
      RBS Nde I Polyhistidine region Nhe I
281 AGAAGGAGAT ATACAT ATG CGG GGT TCT CAT CAT CAT CAT CAT CAT GGT ATG GCT AGC ATG ACT GGT GGA
      Met Arg Gly Ser His His His His His His Gly Met Ala Ser Met Thr Gly Gly
      Xpress™ epitope
351 CAG CAA ATG GGT CGG GAT CTG TAC GAC GAT GAC GAT AAG GAT CAT CCC TT ACC AAGGGC
      Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp His Pro Phe Thr
      EK recognition site EK cleavage site T7 reverse priming site
411 GAGCTCAACG ATCCGGCTGC TAACAAAGCC CGAAAGGAAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA
  
```

**FIG 1** Schematic illustration and genetic sequence of the pET100 vector. The insertion site is shaded.

**Transformation of BL21 Star cells.** To transform the plasmids into BL21 Star (DE3) One Shot *E. coli* cells (Invitrogen), the cells were thawed, two  $\mu$ l of purified plasmid was added to the cells, and the suspension was incubated on ice for 30 minutes. Next, the cells were heat-shocked by incubation for 30 seconds at 42°C. The transformed cells were then combined with 250  $\mu$ l of SOC medium, and the suspension was incubated for 30 minutes at 37°C with shaking. After incubation, the suspension was combined with 10 ml of 2x-YT containing 20  $\mu$ l of ampicillin and 333  $\mu$ l of 30% sterile glucose solution prior to overnight incubation at 37°C with shaking. Per the manufacturer's recommendations, expression of each recombinant protein was then induced prior to freezing. Specifically, two ml of the culture was added to 200 ml of 2x-YT media containing 400  $\mu$ l ampicillin and 7 ml of 30% sterile glucose solution, and the culture was incubated for two hours at 37°C with shaking. After incubation, 200  $\mu$ l of 1M isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) was added, and the culture was re-incubated for 24-36 hours at 26°C. After incubation, the cell suspension was divided into 50 ml fractions, the cells were pelleted by centrifugation, and the pelleted organisms were frozen at -80°C until used.

**Purification of proteins.** Recombinant proteins were recovered by using a commercial kit (Probond Purification System, Invitrogen). Briefly, the cell pellets were lysed by resuspension in 8 ml of guanidinium lysis buffer followed by rocking at room temperature for 10 minutes. Next, the cell lysate was sonicated (Branson Sonic Power W350, Danbury, CT), and the cellular debris was pelleted by centrifugation prior to transferring the supernatant to a fresh tube. The supernatant was then added to a column containing a nickel-charged affinity resin (Invitrogen) that bound the polyhistidine tag on

the N-terminus of the recombinant proteins. The bound proteins were then washed by passing two four-ml volumes of wash buffer (pH 6.0) over the columns followed by an additional two four-ml volumes of wash buffer (pH 5.3). An additional 6 one-ml volumes of wash buffer (pH 5.3) were then poured over the column prior to eluting the recombinant proteins by passing 5 separate one ml volumes of elution buffer (pH 4.0). The captured proteins were subsequently dialyzed and concentrated by using a centrifugal filter unit (Amicon®, Millipore Sigma, Burlington, MA).

#### **Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).**

SDS-PAGE was performed using standard techniques. Briefly, elution buffer that contained 11.5 µg of protein was combined with an equal volume of sample buffer, boiled in a water bath for 5 minutes, and allowed to cool to room temperature. A 30 µl amount was then loaded into an individual well of a 0.1% SDS-10-20% gradient polyacrylamide gel, and the proteins were separated by electrophoresing at 200 V for approximately one hour (Criterion, Biorad, Hercules, CA). Next, the gel was stained by immersing in buffer containing 0.125% Coomassie brilliant blue and rocking the solution at room temperature overnight. After staining, the gel was destained by using a buffer that contained methanol and glacial acetic acid, and the approximate sizes of proteins were determined using a standardized molecular weight standard (GE Healthcare, Pittsburgh, PA).

**Enzyme linked immunosorbent assay (ELISA).** Individual wells of microtiter plates (Immunolon 2 HB; Thermo Labsystems, Franklin, MA) were coated with purified recombinant protein diluted to one µg/ml in carbonate coating buffer (0.15 M sodium carbonate and 0.04 M sodium bicarbonate adjusted to pH 9.6) and incubated overnight at

4°C. After incubation, the plates were washed three times with phosphate buffered saline (pH 7.2) containing 0.05% tween (PBS-T), blocked by adding 200 µl PBS-T and 1% bovine serum albumin and incubating at room temperature for one hour. The plates were then washed with PBS-T, a 100 µl volume of immune serum diluted 1:400 with PBS-T was added to each well, and the plates were re-incubated at room temperature for one hour. After an additional wash, a 100 µl volume of horseradish peroxidase-conjugated anti-human IgM or IgG antibodies (SeraCare Life Sciences Inc., Milford, MA) diluted 1:15,000 or 1:20,000 with PBS-T, respectively, was added to each well, and the plates were incubated at room temperature for an additional hour. After incubation, the plates were washed and 100 µl of freshly prepared o-phenylenediamine dihydrochloride (OPD) substrate (Sigma-Aldrich, St. Louis, MO) was added to each well prior to incubation at room temperature for 30 minutes. The reaction was then stopped by adding 100 µl of 1N H<sub>2</sub>SO<sub>4</sub>, and the optical density (OD) value at 490 nm was measured by a spectrophotometer (VersaMax; Molecular Devices, Sunnyvale, CA).

Prior to using the ELISA, the optimum concentration of conjugated anti-IgM or anti-IgG antibodies was determined by performing checkerboard analyses with a known positive serum and a normal serum (50). The levels of IgM and IgG reactivity considered significant were determined by averaging the optical density (OD<sub>490</sub>) value from 15 case-defined normal sera and then considering a significant reactivity as a reading greater than or equal to three standard deviations above the mean OD value. The result of the analyses were OD cut-off values that ranged from 0.46 to 1.05 (Table 2). A serially-diluted positive control was also included with each ELISA plate and interassay

variability was minimized by requiring the significant reactivity of the control serum to fall within one dilution during each successive run.

**Statistics.** The relatedness among the ELISAs was evaluated by using Cochran's Q test to compare the positive and negative values. A p value  $\leq 0.05$  was considered significant.

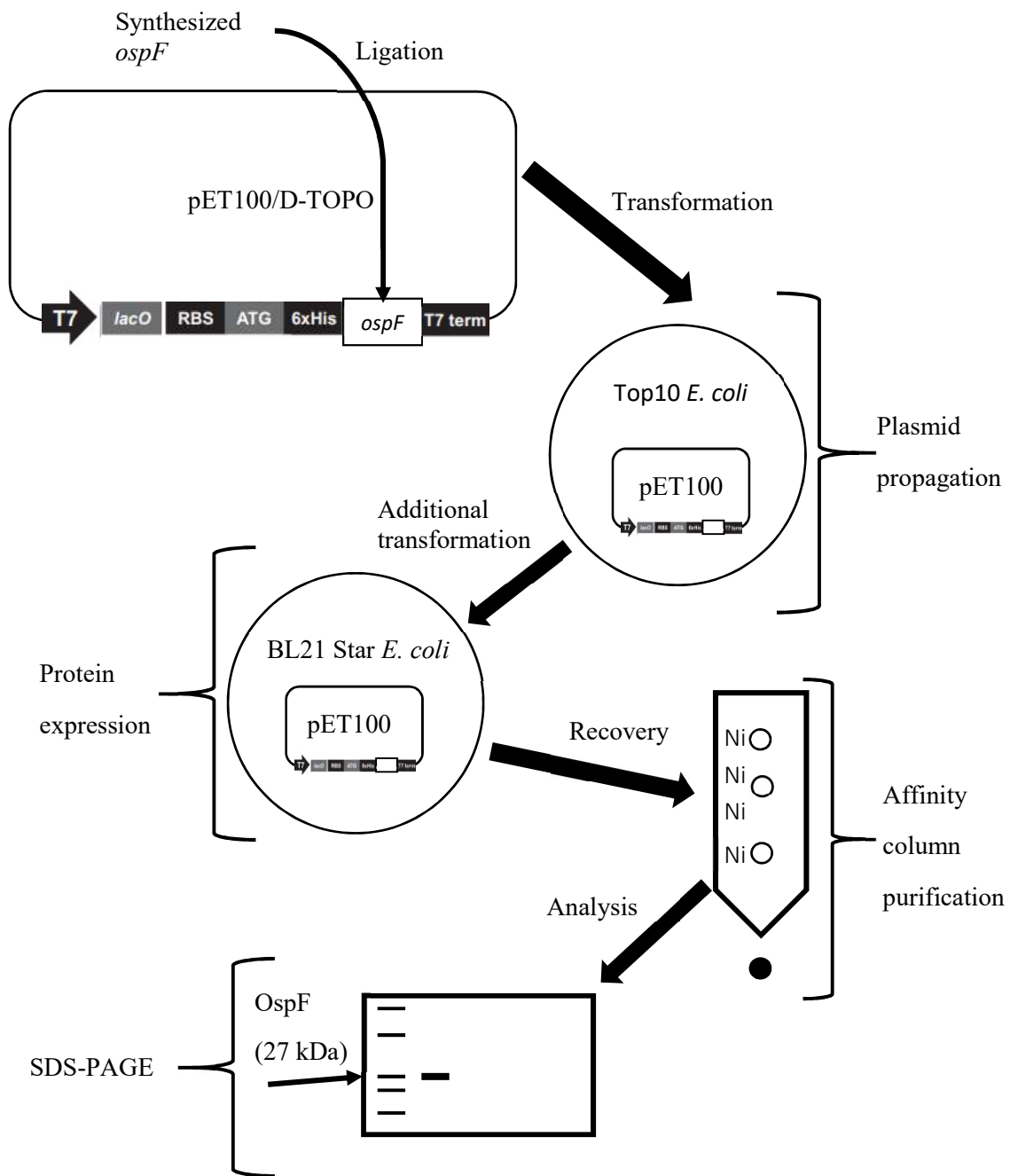
**TABLE 2** Significant optical density (OD) cut-off values for ErpK, ErpL, ErpG or OspF ELISA.

ELISA	Mean optical density (OD) values of normal sera (n=15)		Standard deviation		Significant cut-off value <sup>a</sup>	
	IgM	IgG	IgM	IgG	IgM	IgG
ErpK	0.25	0.40	0.15	0.13	0.70	0.79
ErpL	0.26	0.56	0.11	0.17	0.58	1.05
ErpG	0.21	0.37	0.08	0.13	0.46	0.74
OspF	0.24	0.35	0.14	0.13	0.67	0.74

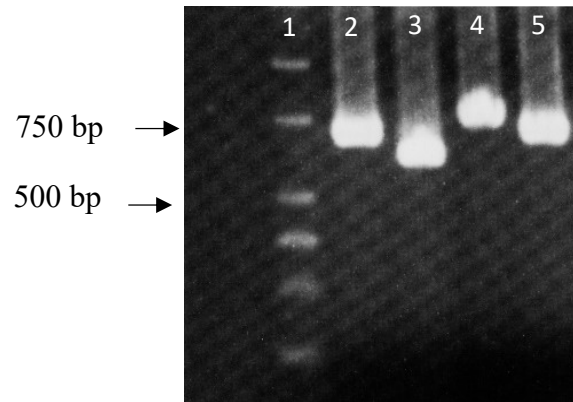
<sup>a</sup> Three standard deviations above the mean OD values of normal sera.

## RESULTS

**Recovery of recombinant (r) proteins.** The cloning strategy was comprised of several distinct steps (Fig. 2), so it was necessary to ensure the validity of the recombinant proteins encoded by each construct by confirming the integrity of the products after each step. Amplification of the genes contained within pUC57 yielded appropriately-sized amplicons (Fig. 3), and the DNA sequences exactly matched the sequence deposited in GenBank after the DNA was purified, inserted into pET100, and transformed into TOP10 *E. coli* (Fig. 4-7). The molecular weight of each recombinant protein recovered after expressing each gene in BL21 Star *E. coli* was also appropriate (Fig. 8). In addition, the purified recombinants contained only small amounts of residual *E. coli* proteins.



**FIG 2** Schematic illustration of cloning strategy.



**FIG 3** Detection of amplified genes by agarose gel electrophoresis. Lane 1 contains 50 bp -10 kbp Perfect DNA Marker, lane 2 contains *ospF* fragment, lane 3 contains *erpG* fragment, lane 4 contains *erpK* fragment, and lane 5 contains *erpL* fragment.

```

Query 1  ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA 60
Sbjct 1  ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA 60
Query 61  ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTCACCGATTTAGAAGGG 120
Sbjct 61  ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTCACCGATTTAGAAGGG 120
Query 121  GCAGTGCAAGATTTAGAAAAGTTCAGAACAAAATGTAAAAAAAACAGAACAAGAGATAAAA 180
Sbjct 121  GCAGTGCAAGATTTAGAAAAGTTCAGAACAAAATGTAAAAAAAACAGAACAAGAGATAAAA 180
Query 181  AAACAAGTTGAAGGATTTTTAGAAAATTCTAGAGACGAAAGATTTGAATAAATTGGATACA 240
Sbjct 181  AAACAAGTTGAAGGATTTTTAGAAAATTCTAGAGACGAAAGATTTGAATAAATTGGATACA 240
Query 241  AAAGAGATTGAAAAACGAATTCAGAATTAAGGAAAAAATAGAAAAATTAGATTCTAAA 300
Sbjct 241  AAAGAGATTGAAAAACGAATTCAGAATTAAGGAAAAAATAGAAAAATTAGATTCTAAA 300
Query 301  AAAAATTCTATTGAAACATATTCTGAGTATGAAGAAAAACTAAAACAAATAAAAGAAAAA 360
Sbjct 301  AAAAATTCTATTGAAACATATTCTGAGTATGAAGAAAAACTAAAACAAATAAAAGAAAAA 360
Query 361  TTGAAAGGAAAGGCAGATCTTGAAGATAAATTAAGGGACTTGAAGATAGCTTAAAAAAG 420
Sbjct 361  TTGAAAGGAAAGGCAGATCTTGAAGATAAATTAAGGGACTTGAAGATAGCTTAAAAAAG 420
Query 421  AAAAAAGAGGAAAGAAAAAAGCTTTAGAAGATGCTAAGAAGAAATTTGAAGAGTTTAAA 480
Sbjct 421  AAAAAAGAGGAAAGAAAAAAGCTTTAGAAGATGCTAAGAAGAAATTTGAAGAGTTTAAA 480
Query 481  GGACAAGTTGGATCCGCGACTGGAGTAACTACCGGGCATAGAGCTGGAAATCAAGGTAGT 540
Sbjct 481  GGACAAGTTGGATCCGCGACTGGAGTAACTACCGGGCATAGAGCTGGAAATCAAGGTAGT 540
Query 541  ATTGGGGCACAAAGCTTGGCAGTGTGCTAATAGTTTGGGGTTGGGTGTAAGTTATTCTAGT 600
Sbjct 541  ATTGGGGCACAAAGCTTGGCAGTGTGCTAATAGTTTGGGGTTGGGTGTAAGTTATTCTAGT 600
Query 601  AGTACTGGTACTGATAGCAATGAATTGGCAAACAAAGTTATAGATGATTCAATTA AAAAAG 660
Sbjct 601  AGTACTGGTACTGATAGCAATGAATTGGCAAACAAAGTTATAGATGATTCAATTA AAAAAG 660
Query 661  ATTGATGAAGAGCTTAAAAATACTATAGAAAATAATGGAGAAGTCAAAAAAGAATAATAA 720
Sbjct 661  ATTGATGAAGAGCTTAAAAATACTATAGAAAATAATGGAGAAGTCAAAAAAGAATAATAA 720

```

**FIG 4** Comparison of recovered *ospF* DNA sequence to the requested sequence. Vertical dashes indicate exact matches.

```

Query 1  ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA 60
Sbjct 1  ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA 60
Query 61  ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTCACCGAAGATTTAAAA 120
Sbjct 61  ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTCACCGAAGATTTAAAA 120
Query 121  CAAAATGTAAAAGAAAAAGTTGAAGGATTTTTAGATAAAGAGTTAATGCAAGGTGACGAT 180
Sbjct 121  CAAAATGTAAAAGAAAAAGTTGAAGGATTTTTAGATAAAGAGTTAATGCAAGGTGACGAT 180
Query 181  CCTAATAACAGTCTGTTTAATCCACCACCAGTATTGCCGGCAAGTCCCACGATAACACA 240
Sbjct 181  CCTAATAACAGTCTGTTTAATCCACCACCAGTATTGCCGGCAAGTCCCACGATAACACA 240
Query 241  CCCGTATTA AAAAGCGGTGCAAGCAAAAAGATGGTGGTCAACAAGAAGGAAAAAGAGAGAAA 300
Sbjct 241  CCCGTATTA AAAAGCGGTGCAAGCAAAAAGATGGTGGTCAACAAGAAGGAAAAAGAGAGAAA 300
Query 301  GAAAAAGAAATTC AAGAATTA AAGGATAAAAATAGATAAAAAGAAAAAAGAATTAGAAGAG 360
Sbjct 301  GAAAAAGAAATTC AAGAATTA AAGGATAAAAATAGATAAAAAGAAAAAAGAATTAGAAGAG 360
Query 361  GCTAGAAAGAAATTTCAAGAATTTAAAGAACAAGTTGAATCTGCAACTGGAGAAAGTACT 420
Sbjct 361  GCTAGAAAGAAATTTCAAGAATTTAAAGAACAAGTTGAATCTGCAACTGGAGAAAGTACT 420
Query 421  GAGAAAGTTAAAAAACAAGGAAATATTGGACAAAAAGCTTTAAAGTATGCTAAAGAAATTG 480
Sbjct 421  GAGAAAGTTAAAAAACAAGGAAATATTGGACAAAAAGCTTTAAAGTATGCTAAAGAAATTG 480
Query 481  GGTGTAATGGAAGTTATTCTGTTAATGATGGTACTAATACTAATGATTTTGTAAAAAAG 540
Sbjct 481  GGTGTAATGGAAGTTATTCTGTTAATGATGGTACTAATACTAATGATTTTGTAAAAAAG 540
Query 541  GTTATAGATGATGCTCTTAAAAATATTGAGGAAGAAGCTTGAAAAGCTAGCAGAGCCTCAA 600
Sbjct 541  GTTATAGATGATGCTCTTAAAAATATTGAGGAAGAAGCTTGAAAAGCTAGCAGAGCCTCAA 600
Query 601  AATATAGAAGATAAAAAATAATAA 624
Sbjct 601  AATATAGAAGATAAAAAATAATAA 624

```

**FIG 5** Comparison of recovered *erpG* DNA sequence to the requested sequence. Vertical dashes indicate exact matches.

```

Query 1   ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA 60
Sbjct 1   ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA 60
Query 61  ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTCACCGTAAAAAAAAGT 120
Sbjct 61  ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTCACCGTAAAAAAAAGT 120
Query 121 TTAGAACAAAGATTTAAAAGGAAAAAGTTAAGGGATTTTATAGATACAAAAAAGAAGAGTTT 180
Sbjct 121 TTAGAACAAAGATTTAAAAGGAAAAAGTTAAGGGATTTTATAGATACAAAAAAGAAGAGTTT 180
Query 181 TTTGGAGATTTTAAAAAACAGAAAGCAAAAGTTCAACCAAAGGATGAAGAATCTATGCAA 240
Sbjct 181 TTTGGAGATTTTAAAAAACAGAAAGCAAAAGTTCAACCAAAGGATGAAGAATCTATGCAA 240
Query 241 GCTGATGAGCCTCAGGAGCAGGGCGAAGATCAAGTGGTTCAAGGTGTTGCTGAAGATCAA 300
Sbjct 241 GCTGATGAGCCTCAGGAGCAGGGCGAAGATCAAGTGGTTCAAGGTGTTGCTGAAGATCAA 300
Query 301 AAATTAAGAAGAAATTGAACAAAAATAAAAGAATTAAGGATAAGATAGAGAAGTCG 360
Sbjct 301 AAATTAAGAAGAAATTGAACAAAAATAAAAGAATTAAGGATAAGATAGAGAAGTCG 360
Query 361 GATCCTAAAAGTGTCTCTTAAAACATATTCTGATTATGAAAAGGAAATAGAAGAATTA 420
Sbjct 361 GATCCTAAAAGTGTCTCTTAAAACATATTCTGATTATGAAAAGGAAATAGAAGAATTA 420
Query 421 AAGGAAAAACTAAAAGATAAAGAAAAGTTGAAAAAGAATTAGAAATCCTTGAGAAAGCT 480
Sbjct 421 AAGGAAAAACTAAAAGATAAAGAAAAGTTGAAAAAGAATTAGAAATCCTTGAGAAAGCT 480
Query 481 TTAATGAAAAAATAGAGAAAAGAAAAAAGAATTAGAAGAGTCTCAAAAAAATTTGAA 540
Sbjct 481 TTAATGAAAAAATAGAGAAAAGAAAAAAGAATTAGAAGAGTCTCAAAAAAATTTGAA 540
Query 541 GAACTTAAGGGTCAAGTTGAATCTGCAATTGGAATAACTGATGGAGAAAGGGCTAAAAAT 600
Sbjct 541 GAACTTAAGGGTCAAGTTGAATCTGCAATTGGAATAACTGATGGAGAAAGGGCTAAAAAT 600
Query 601 CAAGGGAAAAGTTGGAATAGAAGCCTTGCGGCATGCTCGGGGTTTGGGCTTTAAAAATATT 660
Sbjct 601 CAAGGGAAAAGTTGGAATAGAAGCCTTGCGGCATGCTCGGGGTTTGGGCTTTAAAAATATT 660
Query 661 TCTAGTGGTAATAGTACTAGCGATATAGCTAAAGAAATAATAGTAAGCTCGCTTAAAAAG 720
Sbjct 661 TCTAGTGGTAATAGTACTAGCGATATAGCTAAAGAAATAATAGTAAGCTCGCTTAAAAAG 720
Query 721 ATTGAAGAAGAGCTTGAAGAGCTTAAAAAGTTAGAGAAAAGAACTAAAGATTCTAATAAA 780
Sbjct 721 ATTGAAGAAGAGCTTGAAGAGCTTAAAAAGTTAGAGAAAAGAACTAAAGATTCTAATAAA 780
Query 781 AAAGAATAGTAA 792
Sbjct 781 AAAGAATAGTAA 792

```

**FIG 6** Comparison of recovered *erpK* DNA sequence to the requested sequence. Vertical dashes indicate exact matches.

```

Query 1   ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA 60
Sbjct 1   ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA 60

Query 61  ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTACCCCTAAAAAATTC 120
Sbjct 61  ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTACCCCTAAAAAATTC 120

Query 121 GAACAAAATCTAGAAAGTTCAGAACAATAATGTAACAAAAACAGAACAGAGATAAAAAA 180
Sbjct 121 GAACAAAATCTAGAAAGTTCAGAACAATAATGTAACAAAAACAGAACAGAGATAAAAAA 180

Query 181 CAAGTTGAAGGATTTTTAGAAATCTAGAGACAAAAGATTTATCTAAATTAGATGAAAA 240
Sbjct 181 CAAGTTGAAGGATTTTTAGAAATCTAGAGACAAAAGATTTATCTAAATTAGATGAAAA 240

Query 241 GATACAAAAGAAATGAAAAACAATTCAGAATTAAGAATAAAATAGAAAAATTAGAT 300
Sbjct 241 GATACAAAAGAAATGAAAAACAATTCAGAATTAAGAATAAAATAGAAAAATTAGAT 300

Query 301 TCTAAAAAACTTCTATTGAAACATATTCTGAGTATGAAGAAAAAATAAACAAAAATAAA 360
Sbjct 301 TCTAAAAAACTTCTATTGAAACATATTCTGAGTATGAAGAAAAAATAAACAAAAATAAA 360

Query 361 GAAAAATTGAAGGAAAAGGACTTGAAGATAAATTTAAGGAGCTTGAAGAGAGTTTAGCA 420
Sbjct 361 GAAAAATTGAAGGAAAAGGACTTGAAGATAAATTTAAGGAGCTTGAAGAGAGTTTAGCA 420

Query 421 AAGAAAAAGGGGGAGAGAAAAAAGCTTTACAAGAGGCCAAACAGAAATTTGAAGAATAT 480
Sbjct 421 AAGAAAAAGGGGGAGAGAAAAAAGCTTTACAAGAGGCCAAACAGAAATTTGAAGAATAT 480

Query 481 AAAAAACAAGTAGATACTTCAACTGGGAAAACTCAAGGCGACAGGTCTAAAAACCGAGGT 540
Sbjct 481 AAAAAACAAGTAGATACTTCAACTGGGAAAACTCAAGGCGACAGGTCTAAAAACCGAGGT 540

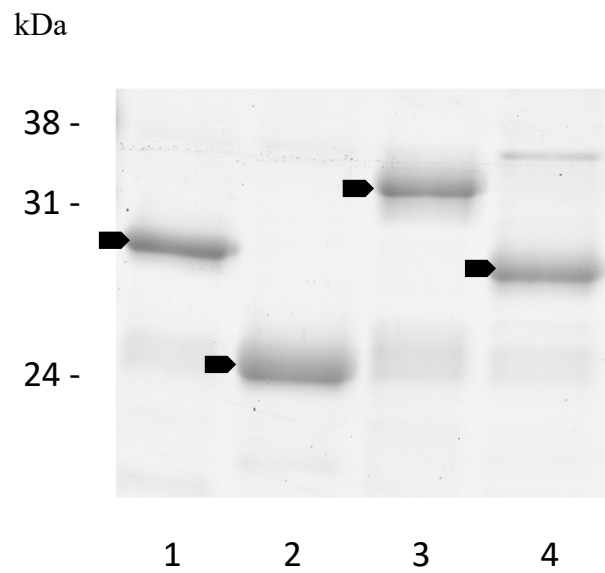
Query 541 GGTGTTGGAGTGCAAGCTTGGCAGTGTCCAATGAATTAGGTTTGGGTGTAAGTTATTCT 600
Sbjct 541 GGTGTTGGAGTGCAAGCTTGGCAGTGTCCAATGAATTAGGTTTGGGTGTAAGTTATTCT 600

Query 601 AATGGCGGAGTGACAACAGCAATACTGATGAATTAGCAAAACAAAGTTATAGATGATTCT 660
Sbjct 601 AATGGCGGAGTGACAACAGCAATACTGATGAATTAGCAAAACAAAGTTATAGATGATTCT 660

Query 661 CTTAAAAAGATTGAAGAAGAACTTAAGGGAAATAGAAGAAGATAAAAAAGAAATAATA 717
Sbjct 661 CTTAAAAAGATTGAAGAAGAACTTAAGGGAAATAGAAGAAGATAAAAAAGAAATAATA 717

```

**FIG 7** Comparison of recovered *erpL* DNA sequence to the requested sequence. Vertical dashes indicate exact matches.



**FIG 8** Detection of recombinant proteins (highlighted by **➡**) by SDS-PAGE. Lane 1 contains OspF, lane 2 contains ErpG, lane 3 contains ErpK, and lane 4 contains ErpL. Other bands are residual *E. coli* proteins.

**Characterization of Lyme disease sera.** A total of 111 Lyme disease sera confirmed by Western blot were used. Of these, 25 (23%) contained *B. burgdorferi* specific IgM antibodies, 22 (20%) contained IgM and IgG antibodies, and 64 (58%) contained IgG antibodies (Table 3). In addition, the clinical complaints included non-specific (e.g. headache, fatigue) constitutional symptoms (n = 23), single or multiple skin lesions (n = 30), Bell's palsy (n = 10), neuropathy/neuritis (n = 2), heart abnormalities (n = 3), arthralgia (n = 18) or arthritis (n = 25).

**TABLE 3** Primary clinical abnormalities of patients with Lyme disease confirmed by CDC Western blot criteria<sup>a</sup>.

Primary clinical complaint (no.)	No. (%) of sera confirmed by:		
	IgM	IgM and IgG	IgG
Constitutional <sup>b</sup> (23)	4 (17)	6 (26)	13 (57)
Erythema migrans <sup>c</sup> (30)	13 (43)	4 (13)	13 (43)
Arthralgia (18)	3 (17)	3 (17)	12 (37)
Bell's Palsy (10)	4 (40)	3 (30)	3 (30)
Cardiac <sup>d</sup> (3)	1 (33)	1 (33)	1 (33)
Neuropathy/Neuritis (2)	0	0	2 (100)
Arthritis (25)	0	5 (20)	20 (80)

<sup>a</sup> IgM= at least two of the 22kDa, 39 kDa, or the 41 kDa proteins  
 IgG= at least five of the 18 kDa, 21 kDa, 28 kDa, 30 kDa, 39 kDa, 41 kDa, 45 kDa, 58 kDa, 66 kDa, or 93 kDa proteins.

<sup>b</sup> Symptoms included fatigue, headache, fever, and myalgia.

<sup>c</sup> Includes patients with single (n=17) or multiple (n=13) EM lesion(s).

<sup>d</sup> Heart complications were bradycardia (n=2), heart block (n=2), asystole (n=1), or atrial fibrillation (n=1).

**Detection of ErpG, ErpK, and ErpL antibodies.** Significant levels of IgM and/or IgG antibodies that bound ErpG, ErpK, or ErpL were detected in the Lyme disease sera, but the response was without apparent pattern. For example, ErpG antibodies were detected in 16 (53%), ErpK was detected in 17 (57%), and ErpL was detected in 11 (37%) sera from patients with EM lesions (Table 4). However, the response failed to expand significantly after the spirochetes disseminated, since similar percentages of the sera from patients with arthralgia or arthritis contained antibodies. In addition, there were no readily-apparent differences when the positive and negative results were compared ( $p = 0.44$ ).

**TABLE 4** Detection of ErpG, ErpK, or ErpL antibodies in Lyme disease sera.

Primary clinical complaint (no.)	No. (%) positive sera								
	ErpG			ErpK			ErpL		
	IgM	IgG	T <sup>a</sup>	IgM	IgG	T	IgM	IgG	T
Constitutional <sup>b</sup> (23)	4 (17)	7 (30)	10 (43)	2 (9)	3 (13)	3 (13)	1 (4)	8 (35)	8 (35)
Erythema migrans <sup>c</sup> (30)	13 (43)	7 (23)	16 (53)	13 (43)	9 (30)	17 (57)	7 (23)	5 (17)	11 (37)
Bell's Palsy (10)	4 (40)	2 (20)	5 (50)	3 (30)	3 (30)	5 (50)	3 (30)	2 (20)	3 (20)
Neuropathy/ Neuritis (2)	0	2 (100)	2 (100)	0	1 (50)	1 (50)	0	1 (50)	1 (50)
Cardiac <sup>d</sup> (3)	0	2 (67)	2 (67)	0	3 (100)	3 (100)	0	2 (67)	2 (67)
Arthralgia (18)	4 (22)	8 (44)	10 (56)	4 (22)	10 (56)	12 (67)	4 (22)	11 (61)	13 (72)
Arthritis (25)	2 (8)	11 (44)	12 (48)	1 (4)	8 (32)	9 (36)	2 (8)	15 (60)	16 (64)

<sup>a</sup> T= Total no. (%) with IgM and/or IgG antibodies.

<sup>b</sup> Symptoms included fatigue, headache, fever, and myalgia.

<sup>c</sup> Includes patients with single (n=17) or multiple (n=13) EM lesion(s).

<sup>d</sup> Heart complications were bradycardia (n=2), heart block (n=2), asystole (n=1), or atrial fibrillation (n=1).

**Detection of OspF antibodies.** Similar to the findings with the Erp ELISAs, significant levels of IgM and/or IgG antibodies were detected in 13 (43%) of patients with likely localized Lyme disease manifesting as single or multiple EM lesions. However, in contrast to the Erp findings, the frequency of detection of OspF antibodies increased coincident with increased duration of infection and severity of symptoms (Table 5). Most notably, significant levels of OspF antibodies were detected in patients with likely joint infection manifesting as arthralgia (78%) or arthritis (84%), and the antibodies were detected in every patient with neuropathy/neuritis or heart complications. The collective findings therefore failed to provide compelling support for the diagnostic utility of the Erps, but confirmed that OspF antibodies were a more predictable immune response that was especially prevalent in patients with disseminated Lyme disease.

**TABLE 5** Detection of OspF antibodies in Lyme disease sera.

Primary complaint (no.)	Likely site of infection	No. (%) positive sera		
		IgM	IgG	Total <sup>a</sup>
Constitutional <sup>b</sup> (23)	Unknown	1 (4)	15 (65)	15 (65)
Erythema migrans <sup>c</sup> (30)	Skin	4 (13)	9 (30)	11 (37)
Bell's Palsy (10)	CNS	3 (30)	4 (40)	5 (50)
Neuropathy/Neuritis (2)		0	2 (100)	2 (100)
Cardiac <sup>d</sup> (3)	Heart	0	3 (100)	3 (100)
Arthralgia (18)	Joints	5 (28)	14 (78)	14 (78)
Arthritis (25)		3 (12)	21 (84)	21 (84)

<sup>a</sup> No. (%) with IgM and/or IgG antibodies detected.

<sup>b</sup> Symptoms included fatigue, headache, fever, and myalgia.

<sup>c</sup> Includes patients with single (n=17) or multiple (n=13) EM lesion(s).

<sup>d</sup> Heart complications were bradycardia (n=2), heart block (n=2), asystole (n=1), or atrial fibrillation (n=1).

## DISCUSSION

Lyme disease is the most widespread tickborne illness in the US and confirming the illness by laboratory tests can be problematic. The most commonly used method is the two-test system comprised of a sensitive ELISA followed by a specific Western blot. However, performing the two-test system is cumbersome and expensive. In response, researchers developed C6 and pepC10 peptide ELISAs that improved the speed and affordability of testing, but the improvements were not sufficient to eliminate the CDC recommendation for the two-test system. Therefore, identifying additional antigens that could be useful for Lyme disease diagnostics remains an important goal. In this study, the OspF, ErpG, ErpK, and ErpL proteins were evaluated to determine whether patients with Lyme disease produced significant levels of antibodies specific for each protein and whether the antibody responses were sufficient to warrant additional studies that more completely evaluate this potential as an important diagnostic marker of infection.

Significant numbers of the Lyme disease sera contained antibodies that bound ErpG, ErpK, and/or ErpL, but the responses were without apparent pattern. In addition, none of the antibody responses were particularly predominant in any stage of infection. For example, approximately 50% of the sera from patients with single or multiple EM lesions contained antibodies against one or more of the Erps, but the response failed to expand significantly thereafter. In fact, only 49-67% of patients with arthralgia or

arthritis produced the antibody response. In addition, there were no significant ( $p = 0.44$ ) differences among the responses detected by using either of the ELISAs, which contradicted previous findings where antibody responses against the Erps varied with duration of infection (37, 38). Furthermore, the findings provided scant support for previous reports (35, 39) that ErpG and ErpK could be differentiated by affinity for glial or epithelial cells, respectively. For example, ErpG antibodies were detected in serum samples from the Lyme disease patients with neuritis/neuropathy, but also detected commonly in patients with skin (EM) or joint (arthralgia/arthritis) infection. Similarly, ErpK antibodies were detected in the sera from patients with a probable infection of the nervous system manifesting as Bell's palsy. These contradictory results are likely due to early literature combining results from different OspF-related proteins and a lack of consistent nomenclature. The collective findings therefore failed to provide much support for the diagnostic utility of the Erps.

In contrast, however, the OspF antibody response clearly expanded coincident with the increasing duration of infection and severity of the clinical abnormalities. For example, while relatively few sera from the patients with EM lesions (37%) contained OspF antibodies, the antibody response clearly predominated as the spirochetes disseminated. Most notably, significant levels of OspF antibodies were detected in most (81%) patients with joint pain or swelling and each patient with cardiac abnormality or neuritis/neuropathy. The results therefore contradicted a previous finding (43) that OspF antibodies were only rarely produced by patients with disseminated Lyme disease and instead confirmed the results from another study (42) that showed that OspF antibodies were reliably (80%) produced by patients with Lyme arthritis. More significantly, the

results also refuted previous studies (44, 45) that suggested the heterogeneous antibody responses due to significant genetic differences within *ospF* would hinder reliable detection of the antibody response, since the seropositive patients resided in diverse locations throughout the 3-state geographic region served by the Gundersen Health System.

Therefore, the findings suggest that the predominant OspF antibody response is targeting a conserved region within OspF and additional studies to discriminate the epitope(s) would be useful. Moreover, the results provide strong evidence that, in contrast to the findings using the Erps, OspF may be a valuable serodiagnostic marker that is especially useful after the Lyme disease spirochetes have disseminated. In addition, even though the antibody response was less frequently detected in the patients with early Lyme disease, a test that detects OspF antibodies may still be valuable for these patients if the response occurs independent of the pepC10 and C6 antibody responses. Curiously, Magnarelli et al. (41) showed in 2002 that almost 90% of early Lyme disease sera were seropositive when the positive OspF antibody responses were combined with those obtained using a pepC10 and C6 ELISA, but the possibility was not formally pursued thereafter. The findings in this study, however, provide added impetus for additional effort to also more completely evaluate this possibility.

Several strengths and a weakness of this study should also be highlighted. As strengths, the intent to recover four recombinant proteins at sufficient volumes to perform ELISAs could have been daunting, but the universal cloning strategy was efficient, and the experience gained with each clone streamlined subsequent efforts. In addition, recovering the recombinant proteins by utilization of a His-tag yielded recombinant

proteins with little contamination from *E. coli* proteins. Moreover, the legitimacy of the immune sera was irrefutable, since antibodies against other proteins were sufficient to satisfy the CDC Lyme disease surveillance criteria. A significant shortcoming, however, was that the recombinant proteins still contained a small amount of *E. coli* proteins that could have confounded precise evaluation of protein concentrations and, in turn, affected the accuracy of the ELISA OD values considered significant. In addition, the specificity of each antibody response was not evaluated by also testing sera from patients with unrelated illnesses that may have potentially contained cross-reactive antibodies. In summary, however, identifying additional *B. burgdorferi* ss proteins that reliably induce antibodies in Lyme disease patients remains an important goal, especially if the response can ultimately be used to eliminate the necessity for confirming the illness by using the cumbersome two-test system. While the results in this study provided little evidence of the utility of the Erps for serodiagnosis, the OspF antibody response was reliably produced, especially in patients with disseminated infection with *B. burgdorferi* ss. Therefore, additional studies that more critically evaluate the utility of detection of OspF antibodies for laboratory confirmation of Lyme disease are especially warranted.

## REFERENCES

1. Steere AC, Malawista SE, Bartenhagen NH, Spieler PN, Newman JH, Rahn DW, Hutchinson GJ, Green J, Snyderman DR, Taylor E. 1984. The clinical spectrum and treatment of Lyme disease. *Yale J Biol Med* 57:453–461.
2. Burgdorfer W. 1984. Discovery of the Lyme disease spirochete and its relation to tick vectors. *Yale J Biol Med* 57:515–520.
3. Burgdorfer W, Barbour A, Hayes S, Benach J, Grunwaldt E, Davis J. 1982. Lyme disease- a tick-borne spirochetosis? *Science* 216:1317–1319.
4. Le Fleche A, Postic D, Girardet K, Peter O. 1997. Characterization of *Borrelia lusitaniae* sp. nov. by 16s ribosomal DNA sequence analysis. *Int J Syst Bacteriol* 47:921-925.
5. Wang G, Van Dam AP, Le Fleche A, Postic D, Peter O, Baranton G, De Boer R, Spanjaard L, Dankert J. 1997. Genetic and phenotypic analysis of *Borrelia valaisiana* sp. nov. (*Borrelia* genomic groups VS116 and M19). *Int J Sys Bacteriol* 47:926-932.
6. Postic D, Ras NM, Lane RS, Hendson M, Baranton G. 1998. Expanded diversity among Californian *Borrelia* isolates and description of *Borrelia bissettii* sp. nov. (formerly *Borrelia* group DN127). *J Clin Microbiol* 36:3497–3504.
7. Shapiro ED. 2014. Lyme disease. *N Engl J Med* 370:1724–1731.
8. Centers for Disease Control & Prevention. 2015. Lyme disease. Retrieved from [https://www.cdc.gov/lyme/signs\\_symptoms/index.html](https://www.cdc.gov/lyme/signs_symptoms/index.html).
9. Salzman BE, Stonehouse A, Studdiford J. 2008. Late diagnosis of early disseminated Lyme disease: perplexing symptoms in a gardener. *J Am Board Fam Med* 21:234-236.
10. Kantamaneni V, Sunder V, Bilal M, Vargo S. 2017. A case of early disseminated neurological Lyme disease followed by atypical cutaneous manifestations. *Case Rep Infect Dis* 2017:6598043.

11. Novak C, Harrison A, Aucott J. 2017. Case report early disseminated Lyme disease with carditis complicated by posttreatment Lyme disease syndrome. *Case Rep Infect Dis* 2017:5847156.
12. Sandholm K, Henningsson AJ, Säve S, Bergström S, Forsberg P, Jonsson N, Ernerudh J, Ekdahl KN. 2014. Early cytokine release in response to live *Borrelia burgdorferi* sensu lato spirochetes is largely complement independent. *PLoS One* 9:e108013.
13. Afari ME, Marmoush F, Rehman MU, Gorski U, Yammine JF. 2016. Lyme carditis: an interesting trip to third-degree heart block and back. *Case Reports Cardiol* 2016:1–3.
14. Kristoferitsch W, Aboulenein-Djamshidian F, Jecel J, Rauschka H, Rainer M, Stanek G, Fischer P. 2018. Secondary dementia due to Lyme neuroborreliosis. *Wien Klin Wochenschr* 130:468–478.
15. Löwhagen GB. 1990. Syphilis: test procedures and therapeutic strategies. *Semin Dermatol* 9:152–159.
16. Lane RS, Lennette ET, Madigan JE. 1990. Interlaboratory and intralaboratory comparisons of indirect immunofluorescence assays for serodiagnosis of Lyme disease. *J Clin Microbiol* 28:1774–1779.
17. Johnson B. 2006. Lyme Disease: Serologic Assays for Antibodies to *Borrelia burgdorferi*, p 493-500. In Detrick B, Hamilton R, Folds J (ed), *Manual of Molecular and Clinical Laboratory Immunology, 7th Edition*. ASM Press, Washington, DC. doi: 10.1128/9781555815905.ch57
18. Moore A, Nelson C, Molins C, Mead P, Schriefer M. 2016. Current guidelines, common clinical pitfalls, and future directions for laboratory diagnosis of Lyme disease, United States. *Emerg Infect Dis* 22:1169-1177.
19. Griffin MO, Fricovsky E, Ceballos G, Villarreal F. 2010. Tetracyclines: a pleiotropic family of compounds with promising therapeutic properties. Review of the literature. *Am J Physiol Cell Physiol* 299:539–548.
20. Engstrom SM, Shoop E, Johnson RC. 1995. Immunoblot interpretation criteria for serodiagnosis of Lyme disease. *J Clin Microbiol* 33:419-427.
21. Dressler F, Whalen JA, Reinhardt BN, Steere AC. 1993. Western blotting in the serodiagnosis of Lyme disease. *J Infect Dis* 167:392–400.

22. Mathiesen MJ, Christiansen M, Hansen K, Holm A, Asbrink E, Theisen M. 1998. Peptide-based OspC enzyme-linked immunosorbent assay for serodiagnosis of Lyme borreliosis. *J Clin Microbiol* 36:3474–3479.
23. Zhang JR, Norris SJ. 1998. Genetic variation of the *Borrelia burgdorferi* gene *vl*sE involves cassette-specific, segmental gene conversion. *Infect Immun* 66:3698–3704.
24. Babb K, El-Hage N, Miller JC, Carroll JA, Stevenson B. 2001. Distinct regulatory pathways control expression of *Borrelia burgdorferi* infection-associated OspC and Erp surface proteins. *Infect Immun* 69:4146–4153.
25. Jobe DA, Lovrich SD, Asp KE, Mathiason MA, Albrecht SE, Schell RF, Callister SM. 2008. Significantly improved accuracy of diagnosis of early Lyme disease by peptide enzyme-linked immunosorbent assay based on the borreliacidal antibody epitope of *Borrelia burgdorferi* OspC. *Clin Vaccine Immunol* 15:981–985.
26. Liang FT, Yan J, Mbow ML, Sviat SL, Gilmore RD, Mamula M, Fikrig E. 2004. *Borrelia burgdorferi* changes its surface antigenic expression in response to host immune responses. *Infect Immun* 72:5759–5767.
27. Fikrig E, Barthold SW, Chen M, Grewal IS, Craft J, Flavell RA. 1996. Protective antibodies in murine Lyme disease arise independently of CD40 ligand. *J Immunol* 157:1–3.
28. Mathiesen MJ, Hansen K, Axelsen N, Halkier-Sørensen L, Theisen M. 1996. Analysis of the human antibody response to outer surface protein C (OspC) of *Borrelia burgdorferi sensu stricto*, *B. garinii*, and *B. afzelii*. *Med Microbiol Immunol* 185:121–129.
29. Marques AR. 2015. Laboratory diagnosis of Lyme disease: advances and challenges. *Infect Dis Clin North Am* 29:295–307.
30. Liang FT, Steere AC, Marques AR, B Johnson BJ, Miller JN, Philipp MT. 1999. Sensitive and specific serodiagnosis of Lyme disease by enzyme-linked immunosorbent assay with a peptide based on an immunodominant conserved region of *Borrelia burgdorferi* VlsE. *J Clin Microbiol* 37:3990-3996.
31. Stevenson B, Tilly K, Rosa PA. 1996. A family of genes located on four separate 32-kilobase circular plasmids in *Borrelia burgdorferi* B31. *J Bacteriol* 178:3508–3516.

32. El-Hage N, Babb K, Carroll JA, Lindstrom N, Fischer ER, Miller JC, Gilmore RD, Lamine Mbow M, Stevenson B. 2001. Surface exposure and protease insensitivity of *Borrelia burgdorferi* Erp (OspEF-related) lipoproteins. *Microbiology* 147:821–830.
33. Akins DR, Caimano MJ, Yang X, Cerna F, Norgard M V, Radolf JD. 1999. Molecular and evolutionary analysis of *Borrelia burgdorferi* 297 circular plasmid-encoded lipoproteins with OspE- and OspF-like leader peptides. *Infect Immun* 67:1526–1532.
34. Caimano MJ, Yang X, Popova TG, Clawson ML, Akins DR, Norgard M V, Radolf JD. 2000. Molecular and evolutionary characterization of the cp32/18 family of supercoiled plasmids in *Borrelia burgdorferi* 297. *Infect Imm* 68:1574-1586.
35. Lin YP, Bhowmick R, Coburn J, Leong JM. 2015. Host cell heparan sulfate glycosaminoglycans are ligands for OspF-related proteins of the Lyme disease spirochete. *Cell Microbiol* 17:1464–1476.
36. Brissette CA, Verma A, Bowman A, Cooley AE, Stevenson B. 2009. The *Borrelia burgdorferi* outer-surface protein ErpX binds mammalian laminin. *Microbiology* 155:863–872.
37. McDowell J V, Sung SY, Price G, Marconi RT. 2001. Demonstration of the genetic stability and temporal expression of select members of the Lyme disease spirochete OspF protein family during infection in mice. *Infect Immun* 69:4831–4838.
38. Hefty PS, Jolliff SE, Caimano MJ, Wikel SK, Radolf JD, Akins DR. 2001. Regulation of OspE-related, OspF-related, and Elp lipoproteins of *Borrelia burgdorferi* strain 297 by mammalian host-specific signals. *Infect Immun* 69:3618–3627.
39. Antonara S, Chafel RM, LaFrance M, Coburn J. 2007. *Borrelia burgdorferi* adhesins identified using in vivo phage display. *Mol Microbiol* 66:262–276.
40. Brissette CA, Cooley AE, Burns LH, Riley SP, Verma A, Woodman ME, Bykowski T, Stevenson B. 2008. Lyme borreliosis spirochete Erp proteins, their known host ligands, and potential roles in mammalian infection. *Int J Med Microbiol* 298:257–267.

41. Magnarelli L, Lawrenz M, Fikrig E, Norris SJ. 2002. Comparative reactivity of human sera to recombinant VlsE and other *Borrelia burgdorferi* antigens in class-specific enzyme-linked immunosorbent assays for Lyme borreliosis. *J Med Microbiol* 51: 649-655.
42. Akin E, McHugh GL, Flavell RA, Fikrig E, Steere AC. 1999. The immunoglobulin (IgG) antibody response to OspA and OspB correlates with severe and prolonged Lyme arthritis and the IgG response to P35 correlates with mild and brief arthritis. *Infect Immun* 67:173–181.
43. Hefty PS, Brooks CS, Jett AM, White GL, Wikel SK, Kennedy RC, Akins DR. 2002. OspE-related, OspF-related, and Elp lipoproteins are immunogenic in baboons experimentally infected with *Borrelia burgdorferi* and in human Lyme disease patients. *J Clin Microbiol* 40:4256-4265.
44. Miller JC, El-Hage N, Bab K, Stevenson B. 2000. *Borrelia burgdorferi* B31 Erp proteins that are dominant immunoblot antigens of animals infected with isolate B31 are recognized by only a subset of human Lyme disease patient sera. *J Clin Microbiol* 38:1569-1574.
45. Marconi RT, Sung SY, Hughes CA, Carlyon JA. 1996. Molecular and evolutionary analyses of a variable series of genes in *Borrelia burgdorferi* that are related to *ospE* and *ospF*, constitute a gene family, and share a common upstream homology box. *J Bacteriol* 178:5615–5626.
46. Wagner B, Freer H, Rollins A, Garcia-Tapia D, Erb HN, Earnhart C, Marconi R, Meeus P. 2012. Antibodies to *Borrelia burgdorferi* OspA, OspC, OspF, and C6 antigens as markers for early and late infection in dogs. *Clin Vaccine Immunol* 19:527-535.
47. Centers for Disease Control and Prevention. 1995. Recommendations for test performance and interpretation from the second national conference on serologic diagnosis of Lyme Disease. *Morb Mortal Wkly Rep* 44:590-591.
48. Picken RN, Strle F, Picken MM, Ruzic-Sabljić E, Maraspin V, Lotric-Furlan S, Cimperman J. 1998. Identification of three species of *Borrelia burgdorferi* sensu lato (*B. burgdorferi* sensu stricto, *B. garinii*, and *B. afzelii*) among isolates from acrodermatitis chronica atrophicans lesions. *J Invest Dermatol* 110:211–214.
49. Setubal JC, Reis M, Matsunaga J, Haake DA. 2006. Lipoprotein computational prediction in spirochaetal genomes. *Microbiology* 152:113–121.

50. ThermoFisher Scientific. 2017. ELISA development and optimization. Retrieved from <https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/elisa-development-optimization.html>

APPENDIX A

TRUNCATED SEQUENCES OF OSPF, ERPG, ERPK, AND ERPL

OspF: Genbank # HM756745.1

ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGA  
CTGGTGGACAGCAAATGGGTCGGGATCTGTACGACGATGACGATA  
AGGATCATCCCTTACCgatttagaaggggcagtgaagatttagaaagttcagaacaaaatgta  
aaaaaacagaacaagagataaaaaacaagtgaaggatttttagaaattctagagacgaaagattgaataaa  
ttggatacaaaaagagattgaaaaacgaattcaagaattaaaggaaaaatagaaaaattagattctaaaaaactt  
ctattgaaacatattctgagtatgaagaaaaactaaaacaataaaagaaaaattgaaaggaaaggcagatcttg  
aagataaattaaaggacttgaagatagcttaaaaaagaaaaagaggaaagaaaaaagctttaagaatgct  
aagaagaaattgaagatttaaaggacaagttggactccgcgactggagtaactaccgggcatagagctggaa  
atcaaggtagtattggggcacaagcttggcagtgctaatagttgggggtgggtgtaagttattctagtagtact  
ggtactgatagcaatgaattggcaacaaagttagatgattcaattaaagattgatgaagagcttaaaaatac  
tatagaaaataatggagaagtaaaaaagaataaTAA

ErpG: Genbank # AE001576.1

ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGG  
ACAGCAAATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTT  
ACCgaagatttaaacaaaatgtaaaagaaaaagttgaaggatttttagataaaagagttaatgcaaggtgacgatcctaataac  
agtctgttaatccaccaccagattgccggcaagttcccacgataacacaccgtattaaagcgggcaagcaaaagatggg  
gtcaacaagaaggaaaaagaagagaaaaagaaattcaagaattaaaggataaaatagataaaagaaaaaaagaattaga  
agaggctagaagaaattcaagaattaaagaacaagttgaatctgcaactggagaaagtactgagaaagttaaaaacaagg  
aaatattggacaaaaagcttaaaagtatgctaaagaattgggtgtaaatggaagttattctgttaatgatggtactaataactaatgatt  
tgtaaaaaaggttatagatgatgctcttaaaaatattgaggaagaacttgaaaagctagcagagcctcaaatatagaagataaaa  
aataaTAA

ErpK: Genbank # AE001578.1

ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGG  
ACAGCAAATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTT  
ACCgtaaaaaaaagtttagaacaagatttaaaaggaaaaagtttaaggatttttagatacaaaaaagaagagttttggagattt  
taaaaaaccagaagcaaaagtcaaccaaaggatgaagaatctatgcaagctgatgagcctcaggagcaggcgaagatcaa  
gtggtcaaggtgtgctgaagatcaaaaattaaagaagaattgaacaaaaataaaagaattaaaggataagatagagaagt  
cggatcctaaaagtgttcttaaaacatattctgattatgaaaggaaatagaagaattaaaggaaaaactaaaagataaagaaa  
agtttgaaaaagaattagaatccttgagaaaagctttaaagtaaaatagagaaaaagaaaaaaagaattaaagagctcaaaa  
aaaattgaagaacttaagggtcaagttgaatctgcaattggaataactgatggagaaagggttaaaaatcaagggaagttgga  
atagaagccttgccgatgctcggggttgggcttaaaaatattctagtggaatagtagcgcgatagctaaagaaataatag  
taagctcgttaaaaagattgaagaagagcttgaagagcttaaaaagtttagagaaagaatctaaagattctaataaaaaagaatag  
TAA

ErpL:

ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGG  
ACAGCAAATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTC  
ACCctaaaaaattcagaacaaaatctagaaaagtcagaacaaaatgtaaaaaaacagaacaagagataaaaaacaagtg  
aaggattttagaattctagagacaaaagatttatctaaattagatgaaaaagatacaaaagaaattgaaaaacaattcaagaatt  
aaagaataaaatagaaaattagattctaaaaaacttctattgaaacatattctgagtatgaagaaaaataaacaataaaaga  
aaaattgaaaggaaaaggacttgaagataaatttaaggagcttgaagagagtttagcaaagaaaaagggggagagaaaaaa  
gctttacaagaggccaacagaaatttgaagaatataaaaaacaagtagatacttcaactgggaaaactcaaggcgacaggtct  
aaaaaccgaggtgggtggtgagtgcaagcttggcagtggtgccaatgaattaggttgggtgtaagtatttctaatggcggcagtg  
caacagcaatactgatgaattagcaacaagttatagatgattcttataaaagattgaagaagaacttaagggatagaagaag  
ataaaaaagaataaTAA

Green= Vector sequence

Blue= Overhang added for insertion into vector

Red= Repeated stop codon

APPENDIX B

VECTOR pET100 /D-TOPO SEQUENCE DATA

1 caaggagatg ggcgccaaca gtccccgggc cacggggcct gccaccatac ccacgccgaa  
61 acaagcgctc atgagcccga agtggcgagc cccgatcttcc ccatcgggtga tgtcggcgat  
121 atagggcgcca gcaaccgcac ctgtggcgcc ggtgatgccg gccacgatgc gtcggcgta  
181 gaggatcgag atctcgatcc cgcgaaatta atacgactca ctatagggga attgtgagcg  
241 gataacaatt cccctctaga aataattttg tttacttta agaaggagat atacatatgc  
301 ggggttctca tcatcatcat catcatggta tggctagcat gactgggtga cagcaaatgg  
361 gtcgggatct gtacgacgat gacgataagg atcatccctt caccaagggc gagctcaacg  
421 atccggctgc taacaaagcc cgaaaggaag ctgagttggc tgctgccacc gctgagcaat  
481 aactagcata accccttggg gcctctaaac gggctctgag gagtttttg ctgaaaggag  
541 gaactatata cggatatccc gcaagaggcc cggcagtacc ggcataacca agcctatgcc  
601 tacagcatcc aggggtgacgg tgccgaggat gacgatgagc gcattgttag atttcataca  
661 cgggtgcctga ctgcgttagc aatttaactg tgataaacta cgcattaaa gcttatcgat  
721 gataagctgt caaacatgag aattaattct tgaagacgaa agggcctcgt gatacgccta  
781 tttttatagg ttaatgtcat gataataatg gtttcttaga cgtcaggtgg cacttttcgg  
841 ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg  
901 ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt  
961 attcaacatt tccgtgtcgc ctttattccc ttttttgagg cattttgcct tcctgttttt  
1021 gctcaccagc aaacgctggg gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg  
1081 ggttacatcg aactggatct caacagcggg aagatccttg agagttttcg ccccgaagaa  
1141 cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgaggatatt atcccggtt  
1201 gacgccggggc aagagcaact cggctgccgc atacactatt ctcagaatga cttgggtgag  
1261 tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt  
1321 gctgccataa ccatgagtga taacactgcg gccaaactac ttctgacaac gatcggagga  
1381 ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg ccttgatcgt  
1441 tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgctgca  
1501 gcaatggcaa caacgcttgcg caaactatta actggcgaac tacttactct agcttcccgg  
1561 caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc  
1621 cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggg  
1681 atcattgcag cactggggcc agatggtaag cctcccgtta tcgtagttat ctacacgagc  
1741 gggatgcagg caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg  
1801 attaaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat tgatttaaaa  
1861 cttcattttt aatttaaaaag gatctaggtg aagatccttt ttgataatct catgacaaa  
1921 atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga  
1981 tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg  
2041 ctaccagcgg tggtttgttt gccggatcaa gagctaccaa ctctttttcc gaaggtaact  
2101 ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta gttaggccac  
2161 cacttcaaga actctgtagc accgcctaca tacctcgtc tgctaactct gttaccagtg  
2221 gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg  
2281 gataaggcgc agcggctcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga  
2341 acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc  
2401 gaagggagaa aggcggacag gtatccggta agcggcaggg tcggaacagg agagcgcacg  
2461 agggagcttc cagggggaaa cgctggtat ctttatagtc ctgtcgggtt tcgccacctc  
2521 tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc  
2581 agcaacgcgg cttttttacg gttcctggcc ttttgctggc cttttgctca catgttcttt  
2641 cctgcgttat cccctgattc tgtggataac cgtattaccg cttttgagtg agctgatacc  
2701 gctcgcgcga gccgaacgac cgagcgcagc gactcagtga gcgaggaagc ggaagagcgc  
2761 ctgatgcggg attttctcct tacgcactgt tgccgtattt cacaccgcaa tgggtgactc  
2821 tcagtacaat ctgctctgat gccgcatagt taagccagta tacactccgc tacctctacg  
2881 tgactgggtc atggctgcgc cccgacacc gccaacacc gctgacgcgc cctgacgggc  
2941 ttgtctgctc ccggcatccg cttacagaca agctgtgacc gtctccggga gctgcatgtg  
3001 tcagaggttt tcaccgtcat caccgaaacg cgcgaggcag ctgcggtaaa gctcatcagc  
3061 gtggctcgtga agegattcac agatgtctgc ctgttcatec gcgtccagct cgttgagttt  
3121 ctccagaagc gttaatgtct ggcttctgat aaagcgggcc atgttaaggg cggttttttc  
3181 ctgtttggte actgatgcct ccgtgtaagg gggatttctg ttcatggggg taatgatacc  
3241 gatgaaacga gagaggatgc tcacgatacg ggttactgat gatgaacatg cccggttact  
3301 ggaacgttgt gagggtaaac aactggcggg atggatgagg cgggaccaga gaaaaatcac

3361 tcaggggtcaa tgccagcgct tcgttaatac agatgtaggt gttccacagg gtagccagca  
3421 gcatcctgcg atgcagatcc ggaacataat ggtgcagggc gctgacttcc gcgtttccag  
3481 acttttacgaa acacggaaac cgaagaccat tcatgttggt gctcaggctcg cagacgtttt  
3541 gcagcagcag tcgcttcacg ttcgctcgcg tatcggatgat tcattctgct aaccagtaag  
3601 gcaaccccg cagcctagcc gggtcctcaa cgacaggagc acgatcatgc gcacccgtgg  
3661 ccaggaccca acgctgcccg agatgcgccc cgtgcggctg ctggagatgg cggacgcgat  
3721 ggatatgttc tgccaagggt tggtttgccg attcacagtt ctccgcaaga attgattggc  
3781 tccaattctt ggagtgggtga atccgttagc gaggtgccgc cggcttccat tcaggtcgag  
3841 gtggcccggc tccatgcacc gcgacgcaac gcggggaggc agacaaggta tagggcggcg  
3901 cctacaatcc atgccaaccc gttccatgtg ctgcggaggc cggcataaat cgccgtgacg  
3961 atcagcggtc caatgatcga agtttaggctg gtaagagccg cgagcgcgat ttgaagctgt  
4021 ccctgatggc cgtcatctac ctgcctggac agcatggcct gcaacgcggg catcccgatg  
4081 ccgcccgaag cgagaagaat cataatgggg aaggccatcc agcctcgcgt gcgaacgccc  
4141 agcaagacgt agcccagcgc gtcggcggcc atgcccggca taatggcctg cttctcggcg  
4201 aaacgtttgg tggcgggacc agtgacgaag gcttgagcga gggcgtgcaa gattccgaat  
4261 accgcaagcg acaggccgat catcgtcgcg ctccagcga agcggctctc gccgaaaatg  
4321 acccagagcg ctgcccggac ctgtcctacg agttgcatga taaagaagac agtcataagt  
4381 gcggcgacga tagtcatgcc ccgcgcccac cgggaaggagc tgactggggt gaaggctctc  
4441 aagggcatcg gtcgagatcc cgggtgcctaa tgagtgagct aacttacatt aattgcgttg  
4501 cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc  
4561 caacgcgcgg ggagaggcgg tttgcgtatt gggcggcagg gtgggttttc ttttaccag  
4621 tgagacgggc aacagctgat tgcccttcac cgcctggccc tgagagagtt gcagcaagcg  
4681 gtccacgctg gtttgcccca gcaggcgaaa atcctgtttg atgggtggtta acggcgggat  
4741 ataacatgag ctgtcttcgg tatcgtcgta tcccactacc gagatatccg caccaacgcg  
4801 cagcccggac tcggtaatgg cgcgcattgc gccacgcgc atctgatcgt tggcaaccag  
4861 catcgcagtg ggaacgatgc cctcattcag catttgcatg gtttggtgaa aaccggacat  
4921 ggcactccag tcgccttccc gttccgctat cggctgaatt tgattgagag tgagatattt  
4981 atgccagcca gccagacgca gacgcgccga gacagaactt aatgggcccg ctaacagcgc  
5041 gatttgctgg tgacceaatg cgaccagatg ctccacgccc agtcgcgtac cgtcttcag  
5101 ggagaaaata atactgttga tgggtgtctg gtcagagaca tcaagaaata acgcccgaac  
5161 attagtgcag gcagcttcca cagcaatggc atcctgggtca tccagcggat agttaatgat  
5221 cagcccactg acgcgcttgcg cgagaagatt gtgcaccgcc gctttacagg cttcgacgcc  
5281 gcttcgcttct accatcgaca ccaccagcgt ggcacccagt tgatcggcgc gagatttaat  
5341 cgccgcgaca atttgcgacg gcgcgtgcag ggccagactg gaggtggcaa cgccaatcag  
5401 caacgactgt ttgcccgcca gttgttgtgc cacgcgggtg ggaatgtaat tcagctccgc  
5461 catcgcgctt tccacttttt cccgcgtttt cgcagaaacg tggctggcct ggttaccac  
5521 gcgggaaacg gtctgataag agacaccggc atactctgcg acatcgtata acgttactgg  
5581 tttcacatcc accaccctga attgactctc ttccggggcg tatcatgcca taccgcgaaa  
5641 ggttttgccg cattcgatgg tgtccgggat ctgcagcctc tcccttatgc gactcctgca  
5701 ttaggaagca gccagtagt aggttgaggc cgttgagcac cgcccgccga aggaatggtg  
5761 catg