THE ROLE OF RECOMBINATION IN OSPC VARIATION IN LYME DISEASE BORRELIA

C. P. Gibbs, I. Livey and F. Dorner

ABSTRACT

The OspC protein of Lyme disease *Borrelia* (*Borrelia burgdorferi* sensu lato) is highly immunogenic and is a protective antigen, thus an ideal candidate for a subunit vaccine. However, this protein is extremely heterogeneous. This study was undertaken to analyze the variability of the OspC protein at the genetic level. The *ospC* gene was amplified by the polymerase chain reaction (PCR) from 76 *Borrelia burgdorferi* sensu lato strains. PCR products were subjected to restriction fragment length polymorphism (RFLP) analyses and genes from different RFLP types were sequenced. A total of 33 *ospC* RFLP types were identified, and two additional RFLP types were deduced from published *ospC* sequences. Genes from different RFLP types were found to be extremely divergent, while within a given RFLP type, no sequence differences were detected. The majority of amino acid changes are localized to the central, highly variable portion of the mature OspC. Pairwise sequence comparisons indicate a mosaic structure of *ospC*. These results suggest that OspC variation is based on frequent recombination between *ospC* alleles; this genetic exchange is proposed to be mediated by lateral transfer of *ospC* sequences between strains.

KEY WORDS

Lyme disease, Borrelia, ospC, recombination

INTRODUCTION

The OspC protein of Lyme disease (LD) *Borrelia* is highly immunogenic in the natural infection (1-3) and is a protective immunogen in animal models (4, 5). This protein is thus a good candidate for a subunit vaccine and may be of value in the serological diagnosis of LD. OspC is a surface-localized lipoprotein with a molecular weight of approximately 22kD. The protein is encoded by a single gene located on a circular plasmid (6,7); expression of OspC correlates inversely with that of OspA and OspB (8), and is regulated by environmental factors (9,10). Comparisons of OspC proteins from different strains of *Borrelia burgdorferi* (*Bb*) indicate this protein is antigenically extremely heterogeneous (11-13). To investigate the nature and extent of this variation at the genetic level, we have analyzed and compared *ospC* genes from a large collection of LD *Borrelia* strains.

MATERIALS AND METHODS

The *ospC* gene was analyzed from a collection of 76 *Bb* sensu lato strains. The strains were obtained from geographically divergent locations and included isolates from infected humans, rodents and ticks; a list of the strains used and relevant characteristics is given in Livey, et al. (14). Oligonucleotide primers corresponding to the 5' and 3' sequences of the *ospC* gene from strain Orth were used in the polymerase chain reaction (PCR) to amplify the *ospC* gene from each of the 76 strains analyzed. The PCR products were subjected to restriction fragment length polymorphism (RFLP) analysis using the enzymes *Dde*I, *Dpn*II and *Dra*I. For nucleotide sequence analysis, the *ospC* gene

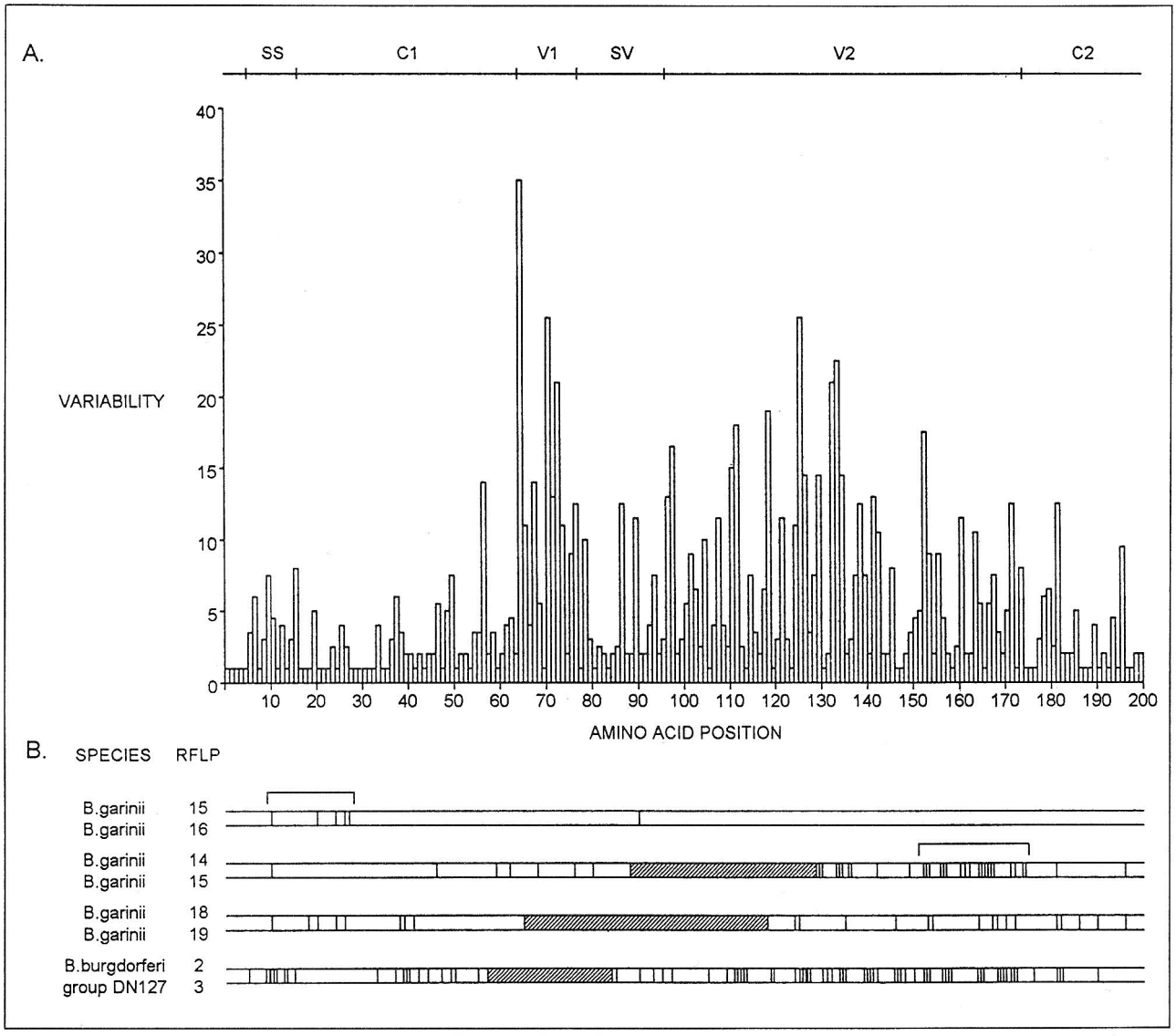

Fig. 1. Variability in the OspC protein and sequence inconsistencies in the ospC gene.
A. Variation in OspC proteins. The sequence variability (ratio of the number of different amino acid residues to the frequency of the most common residue) of the 35 RFLP types is plotted for each amino acid position of the mature protein. Constant and variable domains are indicated.
B. Nucleotide sequence inconsistencies observed in pairwise comparisons of ospC RFLP types. Vertical lines indicate residues differing between the two alleles. Significant stretches of sequence identity are indicated by shading and horizontal bars define regions of significant nucleotide divergence.

was amplified by PCR and the nucleotide sequence of the PCR product determined (14). Amplification of *ospC* genes, RFLP analyses, and nucleotide sequencing have been described (14). Sequence comparisons were performed using the test of Stephens (15).

RESULTS

From analysis of the *ospC* genes of 76 *Bb* sensu lato isolates, 33 different RFLP types were identified (14), and two additional RFLP types could be deduced from published *ospC* sequences (16,17). Forty-two isolates, including at least one from each RFLP type, were selected for further analysis; the nucleotide sequence of the *ospC* gene from each of these strains was determined. Within a given RFLP type, all *ospC* sequences were identical. However, a comparison of *ospC* genes from different RFLP types revealed extreme sequence divergence, with sequence homology ranging from 74.4% to 99.0%.

A comparison of the deduced amino acid sequences of the mature OspC protein from the different *ospC* RFLP types allows the protein to be divided into several domains (Fig. 1A). The amino terminal one-third and carboxy terminal one-eighth form relatively constant domains (C1 and C2), while the central portion of the molecule is extremely variant. This variable region can be further divided into two hypervariable domains (HV1 and HV2) flanking a semi-variable region (SV) centered around a stretch of hydrophobic amino acids. Additionally, close to the carboxy terminus, species-specific motifs (SS) are present.

Pairwise sequence comparisons reveal several sequence inconsistencies, that is, focal regions of extremely high sequence divergence within relatively similar *ospC* genes, or stretches of sequence identity within the variable domains of otherwise dissimilar *ospC* alleles. Some examples are presented in Fig. 1B. For instance, RFLP types 15 and 16, both from *Borrelia garinii* strains, differ in only six nucleotides, but five of these differences are localized in the conserved region near the beginning of the gene. On the other hand, RFLP type 2 (*Bb* sensu stricto) and RFLP type 3 (genogroup DN127) are highly divergent alleles, but the sequence is identical in these two genes in the region spanning the V1 and part of the SV domains. Indeed, short regions of sequence homology are frequently seen, leading to a mosaic nature of *ospC* (Fig. 2).

DISCUSSION

The heterogeneity observed in the OspC protein is characterized by the large number of alleles

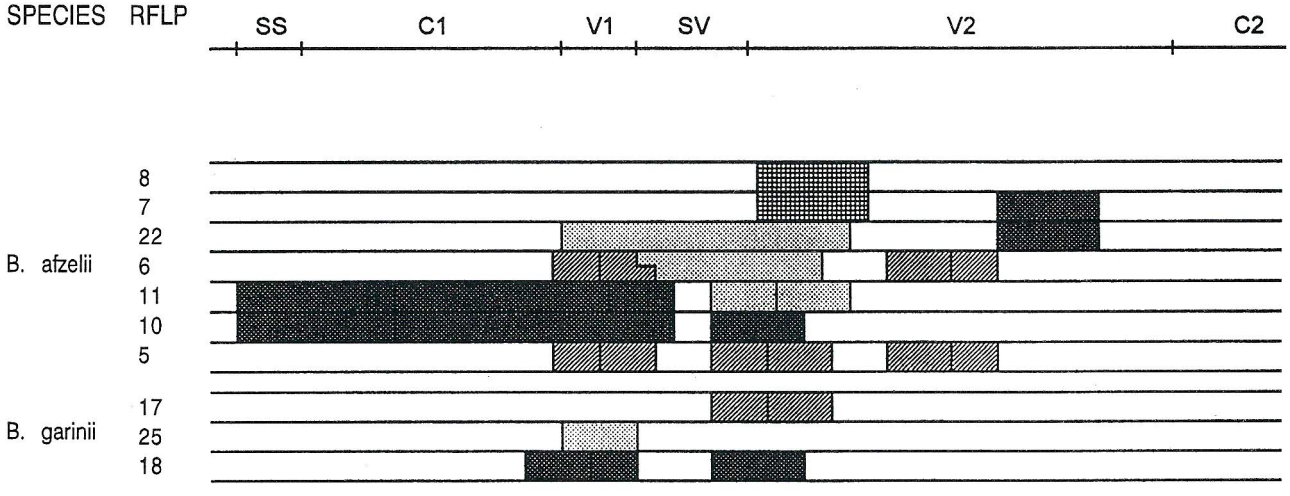

Fig. 2. Mosaic nature of ospC genes. The nucleotide sequences of ospC genes from different RFLP types are compared. Regions of extensive homology between two or more strains within the variable portion of the molecule are indicated by the different shading patterns. Vertical lines indicate nucleotide differences located within homologous stretches.

present in the *Borrelia* population as well as by the extensive sequence divergence between alleles. The amino acid differences are localized primarily in the variable domains in the central portion of the protein, a region containing the most hydrophilic portions of the molecule and thus likely to be immunodominant. The high degree of variability suggests OspC is evolving very rapidly, perhaps subject to intense positive selection. Hypermutability of the *ospC* gene could account for the sequence heterogeneity observed. However, in all cases when *ospC* genes from different strains with the same RFLP type were compared, no nucleotide differences were detected, even among strains from diverse geographic locations. These results would imply that *ospC* genes are relatively stable and that point mutations occur infrequently, a hypothesis that can not be easily reconciled with a mechanism of hypermutability. Any explanation for the molecular mechanism of *ospC* variation must account for both the high variability observed as well as for the lack of genetic intermediates. A recombination-based mechanism for the generation of OspC variability could account for this dichotomy; major sequence changes would result from a single recombination event between different *ospC* alleles. Pairwise sequence comparisons support this hypothesis. Sequence inconsistencies reveal a mosaic structure observed among *ospC* RFLP types (Fig. 2), where scattered, short regions of sequence identity can be found among divergent genes.

The precise mechanism responsible for the recombination among *ospC* genes remains unknown. The *ospC* gene maps to a single locus on a circular plasmid (6,7), and pseudogenes have not been detected (CPG, unpublished). Thus genetic exchange between resident gene copies within a single cell is unlikely to occur. We propose that OspC variation occurs via lateral transfer of genetic material between strains, followed by recombination with the resident *ospC* gene. The location and extent of recombination probably depends upon the degree of sequence homology between the two genes. From analysis of the mosaic structure of the *ospC* genes, it can be seen that genetic transfer probably occurs frequently between strains within a given species, but may also occur between species.

REFERENCES

1. Dressler F, Whalen JA, Reinhardt BN et al. Western blotting in the serodiagnosis of Lyme disease. J Infect Dis 1993; 167: 392-400.

2. Wilske B, Fingerle V, Herzer P et al. Recombinant immunoblot in the serodiagnosis of Lyme Borreliosis. Comparison with indirect immunofluorescence and enzyme-linked immunosorbent assay. Med Microbiol Immunol (Berl) 1993; 182: 255-70.

3. Padula SJ, Dias F, Sampieri A et al. Use of recombinant OspC from *Borrelia burgdorferi* for serodiagnosis of early Lyme disease. J Clin Microbiol 1994; 32: 1733-38.

4. Preac-Mursic V, Wilske B, Patsouris E et al. Active immunization with pC protein of *Borrelia burgdorferi* protects gerbils against *B. burgdorferi* infection. Infection 1992; 20: 342-49.

5. Probert WS, Lefebvre RB. Protection of C3H/HeN mice from challenge with *Borrelia burgdorferi* through active immunization with OspA, OspB, or OspC, but not with OspD or the 83-kilodalton antigen. Infect Immun 1994; 62: 1920-26.

6. Sadziene A, Wilske B, Ferdows MS et al. The cryptic *ospC* gene of *Borrelia burgdorferi* B31 is located on a circular plasmid. Infect Immun 1993; 61: 2192-95.

7. Marconi RT, Samuels DS, Garon CF. Transcriptional analyses and mapping of the *ospC* gene in Lyme disease spirochetes. J Bacteriol 1993; 175: 926-32.

8. Wilske B, Preac-Mursic V, Schierz G et al. Immunochemical analysis of the immune response in late manifestations of Lyme Borreliosis. Zbl Bakt Hyg 1988; 267: 549-58.

9. Schwan TG, Piesman J, Golde WT et al. Induction of an outer surface protein on *Borrelia burgdorferi* during tick feeding. Proc Natl Acad Sci USA 1995; 92: 2909-13.

10. Stevenson B, Schwan TG, and Rosa PA. Temperature-related differential expression of antigens in the Lyme disease spirochete, *Borrelia burgdorferi*. Infect Immun 1995; 63: 4535-39.

11. Theisen M, Frederiksen B, Lebech AM et al. Polymorphism in *ospC* gene of *Borrelia burgdorferi* and immunoreactivity of OspC protein: implications

for taxonomy and for use of OspC protein as a diagnostic antigen. J Clin Microbiol 1993; 31: 2570-76.

12. Wilske B, Preac-Mursic V, Jauris S et al. Immunological and molecular polymorphisms of OspC, an immunodominant major outer surface protein of *Borrelia burgdorferi*. Infect Immun 1993; 61: 2182-91.

13. Stevenson B, Barthold SW. Expression and sequence of outer surface protein C among North American isolates of *Borrelia burgdorferi*. FEMS Microbiol Lett 1994; 124: 367-72.

14. Livey I, Gibbs CP, Schuster R et al. Evidence for lateral transfer and recombination in OspC variation in Lyme disease *Borrelia*. Mol Microbiol 1995; 18: 257-69.

15. Stephens JC. Statistical methods of DNA sequence analysis: Detection of intragenic recombination or gene conversion. Mol Biol Evol 1985; 2: 539-56.

16. Jauris-Heipke S, Fuchs R, Motz M et al. Genetic heterogeneity of the genes coding for the outer surface protein C (OspC) and the flagellin of *Borrelia burgdorferi*. Med Microbiol Immunol (Berl) 1993; 182: 37-50.

17. Padula SJ, Sampieri A, Dias F et al. Molecular characterization and expression of p23 (OspC) from a North American strain of *Borrelia burgdorferi*. Infect Immun 1993; 61: 5097-105.

AUTHORS' ADDRESSES

Carol P. Gibbs, PhD, Immuno AG, Biomedical Research Center, Uferstraße 15, A-2304 Orth an der Donau, Austria
Ian Livey, PhD, same address
Friedrich Dorner, PhD, professor, same address