Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of *Borrelia burgdorferi* outer surface protein C (OspC)

D. Kumaran, S. Eswaramoorthy, J. J. Dunn and S. Swaminathan*

Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA

Correspondence e-mail: swami@bnl.gov

Received 13 September 2000
Accepted 14 November 2000

Single crystals of the outer surface protein C (OspC) from *Borrelia burgdorferi* HB19 have been obtained by the vapor-diffusion method. These crystals belong to space group $P2_1$, with unit-cell parameters $a = 66.218$, $b = 46.113$, $c = 112.079$ Å, $\beta = 99.30^\circ$, and diffract to at least 2.2 Å resolution. Native data have been collected from flash-frozen crystals at the National Synchrotron facility of Brookhaven National Laboratory. There are two dimers per asymmetric unit, related by a non-crystallographic twofold axis and a pseudo-translational symmetry.

1. Introduction

Lyme disease is the most common tick-borne disease in the United States (Barbour & Fish, 1993). Clinical manifestation may involve the skin, nervous system, heart and joints (Asbrink & Hovmark, 1988; Hansen & Lebech, 1992; Steere, 1991; van der Linde *et al.*, 1990). The causative agent is *B. burgdorferi*, a spirochete first isolated from Shelter Island, New York (Burgdorfer *et al.*, 1982). *B. burgdorferi* infection induces a strong humoral response against endoflagellar protein p41 and a protein constituent of the protoplasmic cylinder p93 (Craft *et al.*, 1986), both of which are enveloped within the outer membrane, and some outer-surface lipoproteins (Osps), which are the major components of the membrane (Schwan *et al.*, 1995). Of these, OspA and OspC have been most studied. Experimental vaccines produced against either OspA or OspC are successful in warding off challenges from the same strain (Fikrig *et al.*, 1990; Gilmore *et al.*, 1996; Probert *et al.*, 1997; Simon *et al.*, 1991), but not necessarily from heterologous strains (Fikrig *et al.*, 1992; Golde *et al.*, 1995; Probert *et al.*, 1997).

The expression pattern of OspA and OspC is suggestive of their roles in the infection cycle of the spirochetes (Schwan *et al.*, 1995; Stevenson *et al.*, 1995). In unfed ticks, spirochetes express OspA, but not OspC. However, when the tick starts feeding on mammals, OspC synthesis is induced and OspA repressed. The switch is in part a consequence of the change in temperature; OspC is induced at 305–310 K but not at 297 K, and this upregulation is at the transcriptional and trans-

```
                                          MA\
          aa19                              38----------------------
HB19 MKKNTLSAILMTLFLFISCNNSGKDGNTSANSADESV] [KGPNLTEISKKITESNAVVLAVK 60
B31  MKKNTLSAILMTLFLFISCNNSGKDGNTSANSADESV] [KGPNLTEISKKITDSNAVLLAVK 60
     ************************************   **************:****:****

     ------------------------------------------------------------
HB19 EVETLLTSIDELAK-AIGKKIKNDVSLDNEADHNGSLISGAYLISTLITKKISAIKDSGE 119
B31  EVEALLSSIDEIAAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEG- 119
     ***:**:****:*  ******::: .**.* :*****::*** *****.:*:..:*:.*
     ------------------------------------------------------------
HB19 LKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFES 179
B31  LKEKIDAAKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFES 179
     ** :*: ****** ** *** :*********** :**:******. *****:** *****
     -------------------201
HB19 VKNLSKAAKEMLTNSVKELTSP] [VVAESPKKP 210
B31  VEVLSKAAKEMLANSVKELTSP] [VVAESPKKP 210
     *: *********:*********   *********
```

Figure 1
CLUSTAL alignment of HB19 and B31 OspC proteins. The first 18 residues (underlined) function as a signal sequence for transport, proteolytic processing and lipidation. The constructs used here, aa38-201 and aa38-210, were produced by polymerase chain reaction amplification of genomic DNA with primers that introduced unique restriction sites in front of aa38 and after aa201 or after the end of the *osp*C open reading frame. The 5′ primer provided the initiation codon (M) followed by an A codon; the initiation M residue is removed post-translationally. The 3′ primers had appropriately placed stop codons after residues 201 or 210, respectively. The sequences of the clones were confirmed by DNA sequencing. The asterisks, colons and spaces below the amino-acid sequence represent complete identity, conserved changes and no homology at all, respectively.

lational levels (Ramamoorthy & Philipp, 1998; Tilly *et al.*, 1997). Evidence suggests co-regulation of these two genes at the mRNA level (Jonsson & Bergstrom, 1995). Clearly, to survive in both hosts, spirochetes have evolved mechanisms in sensing the different host environments and in responding accordingly.

The *osp*C gene is located on a 27 kbp circular plasmid and encodes a lipoprotein of about 22–23 kDa (Marconi *et al.*, 1993; Sadziene *et al.*, 1993). The protein is initially synthesized with an 18-amino-acid-long signal sequence which is removed by the cell's signal II peptidase system during processing and lipidation of the N′-terminal Cys residue, amino acid 19. The remainder of the protein is highly polymorphic and this variability extends even to a single *Borrelia* population (strains collected from one geographical area). For example, alleles of OspC collected from a single site on Shelter Island, New York, could be clustered into 19 major groups (*A*–*S*) based on sequence homology (Wang *et al.*, 1999). Sequence variation within a major group is less than 1%, while it is about 19% across the major groups. Variation within a local population is comparable with the variation of similar size samples collected from the entire species. This variability has consequences in the development of OspC-based seradiagnostic antigens and anti-OspC vaccines. Of the 19 major groups, only four groups (*A*, *B*, *I* and *K*) contain invasive clones and cause infections of skin and extracutaneous sites, while the others are non-human pathogens or infect only the skin. In order to develop an effective OspC-based vaccine, it is important to know representative three-dimensional structures of at least a few OspCs, especially from the invasive strains; this information could be useful for rational design of an OspC-based recombinant vaccine. Here, we report the cloning, expression, purification and crystallization of OspC from strain hb19, which is a member of invasive group *I*.

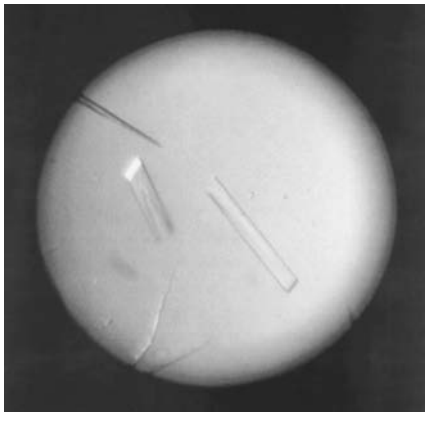

Figure 2
Single crystals of OspC-Hb19, photographed under a polarizing microscope.

2. Materials and methods

Expression of the full-length 219 amino-acid-long HB19 and B31 OspCs is extremely toxic to *Escherichia coli*, causing rapid cell lysis shortly after IPTG induction of a recombinant clone in standard T7-based expression systems (Studier *et al.*, 1990). Therefore, the genes were cloned and expressed successfully without their N′-terminal signal sequences as previously described for expression of OspA (Dunn *et al.*, 1990). Subsequent NMR studies (Huang *et al.*, 1999) on the purified B31 protein showed that the conformation of the first 37 and the last ten residues are very flexible and that the N′- and C′-termini may be in close proximity. Therefore, two additional clones were constructed; one expressing residues 38–210 and the other residues 38–201 (Fig. 1). Details regarding the construction of these clones and purification of the expressed proteins will be published elsewhere. Briefly, the expressed proteins are soluble and can be purified to homogeneity by serial anion (Q Sepharose) and cation (S Sepharose) ion-exchange chromatography. Peak fractions were dialyzed against 10 m*M* sodium phosphate and 50 m*M* sodium chloride pH 6.0. Crystallization trials were carried out with construct 2, amino acids 38–201. Initial crystallization conditions were obtained with the use of Hampton Research Crystal Screens but restricting the trials to pH values between 6 and 8, close to the pI (6.96) of the protein. 4 µl of a droplet containing a 1:1 ratio of protein at a concentration of 6.9 mg ml^{-1} to the precipitant was placed in a microbridge, which was in turn placed inside Linbro plate wells containing appropriate precipitants. The wells were sealed with clear tape and left at room temperature. Initial trials gave needle-shaped crystals with 25% PEG monomethyl ether 550, 10 m*M* zinc sulfate heptahydrate, 100 m*M* MES at pH 6.5. The condition was further improved by varying the concentration of the protein. Long rod-like crystals of dimensions 0.3 × 0.3 × 1.0 mm (Fig. 2) were obtained in about a week using HB19:aa38–201. These crystals diffracted to better than 2.2 Å resolution. The best cryocondition was obtained by adding glycerol to the original mother liquor so that its final concentration was 15%. However, this diluted the other constituents of the mother liquor.

Single-crystal X-ray diffraction data at liquid-nitrogen temperature were collected at beamline X26C of the National Synchrotron Light Source, Brookhaven National laboratory. A Quantum 4 detector was used to collect data with a crystal-to-detector distance of 180 mm with λ = 1.087 Å. Data were reduced with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The crystals belong to the space group $P2_1$, with unit-cell parameters a = 66.218, b = 46.113, c = 112.079 Å, β = 99.30°. Assuming four molecules of 17 684 Da per asymmetric unit, the Matthews coefficient (Matthews, 1968) is 2.36 Å^3 Da^{-1}, corresponding to an estimated solvent content of 49% by volume of the unit cell.

3. Results and discussion

X-ray diffraction data were collected from a flash-frozen crystal at beamline X26C of the NSLS, Brookhaven National Laboratory. A total of 120 095 reflections were collected and reduced to 33 872 unique reflections with no σ cutoff on intensity. The overall $R_{\rm sym}$ was 5.3% (21.5% for the highest resolution shell). The data were 97.3% complete to 2.2 Å resolution, with 88.2% completeness in the outermost shell (2.28–2.20 Å) and an average I/σ of 11.

A self-rotation Patterson function was calculated which showed a non-crystallographic twofold symmetry in the *ac* plane. A stereographic projection of the self-rotation function is shown in Fig. 3. A peak with 56% of the origin peak strength is seen at ψ = 90, φ = 40.5, κ = 180°. A second peak with about 25% of the origin peak also appeared at κ = 12°. A native self-Patterson was calculated and the section corresponding to v = 0

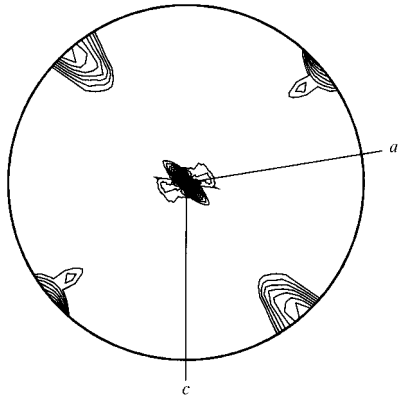

Figure 3
Stereographic projection at κ = 180° of self-rotation function (Rossmann & Blow, 1962) as obtained from the program *ROTRAN* (Craven, 1975). The minimum contour level corresponds to 32% of the maximum, with contour intervals at 4%.

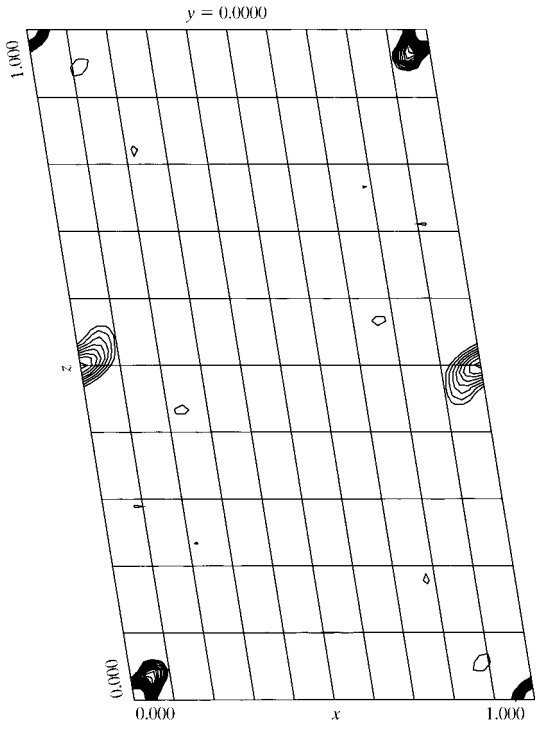

Figure 4
A native Patterson corresponding to section $v = 0$. The peak corresponding to pseudo-translation is seen at $u = 0$, $w = 1/2$. All reflections in the resolution range 40–5 Å were used in the calculations. The map is contoured at intervals of 2% of the origin peak, with the minimum and maximum contour levels of 4% and 60% of the origin peak. At higher resolution the peak strength decreases.

is shown in Fig. 4. There is a pseudo-translation corresponding to $(0, 0, \frac{1}{2})$. The self-Patterson was performed at various resolution cutoffs and the pseudo-translation peak became less strong at higher resolution. Since self-rotation Patterson analysis showed a peak with a very small value of κ ($\sim$12°), the pseudo-translated molecule may involve a slight rotation of the molecule. The HB19 molecule exists as a dimer related by a non-crystallographic twofold axis and there are two dimeric molecules related by a pseudo-translation in the asymmetric unit.

Because of the variability in the OspC sequence, a seradiagnostic antigen or anti-OspC vaccine developed for one *Borrelia* strain will not be effective for all strains. Having the three-dimensional structure of at least one OspC will provide information relating sequence variability and overall structure. This information, when coupled with modeling and biochemical studies, should be useful for developing a vaccine that will be effective against a wide range of different OspC targets. We are attempting to solve the structure of HB19 OspC using the multiple isomorphous replacement with anomalous scattering (MIRAS) method (Blow & Crick, 1959). Recently, we have been able to express selenomethionine-substituted protein in order to determine the structure by the multiple anomalous dispersion (MAD) method (Hendrickson, 1991).

This research was supported by the National Institutes of Health (AI37256) at Brookhaven National Laboratory under contract with the US Department of Energy. DK was supported by the Veterans Administration Medical Center, Pittsburgh. We thank B. Lade for technical assistance and X. Yang for initial help with OspC purification and crystallization trials. We also thank S. Koide for communicating results prior to publication.

References

Asbrink, E. & Hovmark, A. (1988). *Ann. NY Acad. Sci.* **539**, 4–15.
Barbour, A. G. & Fish, D. (1993). *Science*, **260**, 1610–1616.
Blow, D. M. & Crick, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.
Burgdorfer, W., Barbour, A. G., Hayes, S. F., Benach, J. L., Gruwldt, E. & Davis, J. P. (1982). *Science*, **216**, 1317–1319.
Craft, J. E., Fischer, D. K., Shimamoto, G. T. & Steere, A. C. (1986). *J. Clin. Invest.* **78**, 934–939.
Craven, B. M. (1975). *Computer Program ROTRAN for the Determination of Crystal Structures which contain Rigid Body Fragments of Known Structure.* Department of Crystallography, University of Pittsburgh.
Dunn, J. J., Lade, B. N. & Barbour, A. G. (1990). *Protein Expr. Purif.* **1**, 159–168.
Fikrig, E., Barthold, S. W., Kantor, F. S. & Flavell, R. A. (1990). *Science*, **250**, 533–536.
Fikrig, E., Barthold, S. W., Persing, D. H., Sun, X., Kantor, F. S. & Flavell, R. A. (1992). *J. Immunol.* **148**, 2256–2260.
Gilmore, R. D., Kappel, K. J., Dolan, M. C., Burkot, T. R. & Johnson, B. J. (1996). *Infect. Immun.* **64**, 2234–2239.
Golde, W. T., Burkot, T. R., Piesman, J., Dolan, M. C., Capiau, C., Hauser, P., Dequesne, G. & Lobet, Y. (1995). *Vaccine*, **13**, 435–441.
Hansen, K. & Lebech, A. M. (1992). *Brain*, **115**, 399–423.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Huang, X., Link, K., Koide, A., Dunn, J. J., Luft, B. J. & Koide, S. (1999). *J. Biomol. NMR*, **14**, 283–284.
Jonsson, M. & Bergstrom, S. (1995). *Microbiology*, **141**, 1321–1329.
van der Linde, M. R., Crijns, H. J., de Koning, J., Hoogkamp-Korstanje, J. A., de Graaf, J. J., Piers, D. A., van der Galien, A. & Lie, K. I. (1990). *Br. Heart J.* **63**, 162–168.
Marconi, R. T., Samuels, D. S. & Garon, C. F. (1993). *J. Bacteriol.* **175**, 926–932.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Probert, W. S., Crawford, M., Cadiz, R. B. & LeFebvre, R. B. (1997). *J. Infect. Dis.* **175**, 400–405.
Ramamoorthy, R. & Philipp, M. T. (1998). *Infect. Immun.* **66**, 5119–5124.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Sadziene, A., Wilske, B., Ferdows, M. & Barbour, A. (1993). *Infect. Immun.* **61**, 2192–2195.
Schwan, T. G., Piesman, J., Golde, W. T., Dolan, M. C. & Rosa, P. A. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 2909–2913.
Simon, M. M., Schaible, U. E., Kramer, M. D., Eckerskorn, C., Museteanu, C., Muller-Hermelink, K. K. & Wallich, R. (1991). *J. Infect. Dis.* **164**, 123–132.
Steere, A. C. (1991). *Scand. J. Infect. Dis. Suppl.* **77**, 51–54.
Stevenson, B., Schwan, T. G. & Rosa, P. A. (1995). *Infect. Immun.* **63**, 4535–4539.
Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). *Methods Enzymol.* **185**, 60–89.
Tilly, K., Casjens, S., Stevenson, B., Bono, J. L., Samuels, D. S., Hogan, D. & Rosa, P. (1997). *Mol. Microbiol.* **25**, 361–373.
Wang, I.-N., Dykhuizen, D. E., Qui, W., Dunn, J. J., Bosler, E. M. & Luft, B. J. (1999). *Genetics*, **151**, 15–30.