Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of an outer surface lipoprotein BBA64 from the Lyme disease agent *Borrelia burgdorferi* which is critical to ensure infection after a tick bite

Kalvis Brangulis,[a]* Kaspars Tars,[a] Ivars Petrovskis,[a] Andris Kazaks,[a] Renate Ranka[a] and Viesturs Baumanis[a,b]

[a]Latvian Biomedical Research and Study Centre, Ratsupites 1, Riga, LV-1067, Latvia, and [b]University of Latvia, Kronvalda bulv. 4, Riga, LV-1586, Latvia

Correspondence e-mail: kalvis@biomed.lu.lv

Received 11 December 2012
Accepted 27 February 2013

PDB Reference: BBA64, 4aly

Lyme disease is a tick-borne infection caused by the transmission of *Borrelia burgdorferi* from infected *Ixodes* ticks to a mammalian host during the blood meal. Previous studies have shown that the expression of *B. burgdorferi* surface-localized lipoproteins, which include BBA64, is up-regulated during the process of tick feeding. Although the exact function of BBA64 is not known, this lipoprotein is critical for the transmission of the spirochete from the tick salivary glands to the mammalian organism after a tick bite. Since the mechanism of development of the disease and the functions of the surface lipoproteins associated with borreliosis are still poorly understood, the crystal structure of the *B. burgdorferi* outer surface lipoprotein BBA64 was solved at 2.4 Å resolution in order to obtain a better insight into the pathogenesis of *B. burgdorferi* and to promote the discovery of novel potential preventive drugs against Lyme disease. In this study, the crystal structure of BBA64 was also compared with that of the paralogous protein CspA (also referred to as BbCRASP-1, CRASP-1 or BBA68). CspA is the complement regulator-acquiring surface protein-1 of *B. burgdorferi*; its structure is known, but its function apparently differs from that of BBA64. It is demonstrated that unlike the homologous CspA, BBA64 does not form a homodimer. Their differences in function could be explained by divergence in their amino-acid sequences, electrostatic surface potentials and overall tertiary structures. The C-terminal part of BBA64 has a different conformation to that of CspA; the conformation of this region is essential for the proper function of CspA.

1. Introduction

Borrelia burgdorferi, the causative agent of Lyme disease, is transmitted from infected *Ixodes* ticks to a mammalian host organism during the blood meal (Burgdorfer *et al.*, 1982; Steere *et al.*, 2004).

The change in the host organism on transfer from *Ixodes* ticks to mammals and the necessity to proliferate and to resist the immune response of the host has forced *Borrelia* to adapt to changing environments. Studies have shown that *B. burgdorferi* switches the expression of different genes in response to mammalian host-specific signals, pH and temperature shifts or cell-density changes (Angel *et al.*, 2010; Brooks *et al.*, 2003; Carroll *et al.*, 2000; Embers *et al.*, 2004; Indest *et al.*, 1997; Ojaimi *et al.*, 2003; Ramamoorthy & Scholl-Meeker, 2001; Revel *et al.*, 2002; Tokarz *et al.*, 2004).

The *B. burgdorferi* genes coding for the outer surface lipoproteins BBA64 and BBA68 are members of the paralogous gene family Pfam54 located on the 54 kb linear plasmid

(lp54), which is one of the 12 linear plasmids of *B. burgdorferi* (Casjens *et al.*, 2000, 2012; Fraser *et al.*, 1997). It has been suggested that the proteins expressed by the genes residing on lp54 play an important role in the adaptation of *Borrelia* and are associated with borrelial pathogenesis (Hughes *et al.*, 2008; Ojaimi *et al.*, 2003; Tokarz *et al.*, 2004). During a temperature shift *in vitro* from 296 to 308 K, which resembles the temperature change during the transfer of the spirochete from ticks to a warm-blooded animal, and in response to mammalian host-specific signals, the 54 kb linear plasmid had the highest number of differentially expressed genes of all of the borrelial plasmids and the members of the paralogous gene family Pfam54 particularly stood out (Brooks *et al.*, 2003; Ojaimi *et al.*, 2003; Tokarz *et al.*, 2004).

Recent studies have revealed that BBA64 plays a vital role in the transfer of *B. burgdorferi* from *Ixodes* ticks to the host organism after the tick bite (Gilmore *et al.*, 2010). A mutant borrelial strain with an inactivated BBA64 gene is incapable of ensuring the transmission of *B. burgdorferi* and infection is therefore not initiated (Anguita *et al.*, 2000; Gilmore *et al.*, 2007, 2010; Schmit *et al.*, 2011; Patton *et al.*, 2011). However, the exact ligand or receptor for BBA64 is still under investigation.

CspA has 25% sequence identity to BBA64 and is the only member of the Pfam54 proteins for which both structure and function are known. CspA binds complement regulator factor H and factor-H-like protein-1 (FHL-1; Kraiczy, Skerka, Kirschfink, Brade *et al.*, 2001, Wallich *et al.*, 2005) and plays an essential role in decreasing the immune response of the host by inactivating the alternative complement pathway (Kraiczy, Skerka, Kirschfink, Zipfel *et al.*, 2001).

To gain further understanding of exactly how BBA64 ensures the transfer of *B. burgdorferi* from ticks to the mammalian host and to potentially use the protein as a novel drug target to stop the transmission of the spirochete, we have solved the three-dimensional structure of recombinant BBA64 at 2.4 Å resolution and have compared the structure with the previously determined crystal structure of the homologous protein CspA (PDB entry 1w33; Cordes *et al.*, 2004, 2005). The structure of BBA64 revealed an overall fold similar to that of CspA except for the C-terminal α-helix, which has a different conformation compared with that of CspA. In CspA the C-terminal α-helix is involved in the formation of a homodimer and is essential for the proper function of CspA (Cordes *et al.*, 2006; Kraiczy *et al.*, 2009). Although it was originally considered that the cleft between the monomers in CspA could be an ideal place for ligand binding and thus might be conserved in homologous proteins, it is also reasonable to believe that the potential ligand-binding sites may vary between the homologues. This is supported by the facts that dimer formation was not observed in the case of BBA64 and the crystal structure revealed a different orientation of the C-terminal α-helix, which is necessary for dimer formation of CspA, and also revealed the proper binding of complement regulator factor H and FHL-1 in a monomeric state of CspA (Cordes *et al.*, 2006; Kraiczy *et al.*, 2006).

2. Materials and methods

2.1. Cloning and expression of native and SeMet-labelled proteins

The gene for recombinant BBA64 protein was amplified from *B. burgdorferi* strain B31 genomic DNA by PCR so that the signal sequence (amino acids 1–32) of the protein as predicted by *SignalP* 3.0 and *LipoP* 1.0 was excluded (Bendtsen *et al.*, 2004; Juncker *et al.*, 2003). The PCR product with primer-introduced restriction sites for *Nco*I and *Not*I endonucleases was cleaved with the respective enzymes and ligated into the pETm_11 expression vector (EMBL, Heidelberg) encoding a TEV (*Tobacco etch virus*) protease cleavage site followed by a His tag. The obtained plasmid was transformed into *Escherichia coli* strain RR1 and the cells were grown overnight at 310 K on LB agar plates containing kanamycin. Colonies were inoculated into liquid LB medium containing kanamycin at 310 K for a further 24 h. Plasmid DNA was isolated from the resulting culture and DNA sequencing was carried out for all of the obtained clones to ensure that no errors had occurred. For the overexpression of hexahistidine-tagged fusion protein, the plasmid of the correct construct was transformed into *E. coli* BL21 (DE3) cells, which were grown in modified 2×TYP medium [2TY medium containing 16 g l^{-1} bacto-tryptone (Difco), 10 g l^{-1} bacto yeast extract (Difco) and 5 g l^{-1} NaCl, supplemented with 10 mg ml^{-1} kanamycin, 133 m*M* phosphate buffer pH 7.4 and 4 g l^{-1} glucose] with vigorous agitation at 298 K until an OD of 0.8–1.0 was reached and were then induced with 0.2 m*M* IPTG and cultivated for a further 16–20 h.

For the expression of SeMet-labelled protein, the plasmid was transformed into *E. coli* B834 (DE3) cells. The cells were grown in modified 2×TYP medium (supplemented with 133 m*M* phosphate buffer pH 7.4 and 2 g l^{-1} glucose) until an OD of 1.0 was reached and then centrifuged; the pellet was resuspended in 0501 medium (without methionine) (SelenoMet medium base from Athena Enzyme Systems) supplemented with 0502 medium (SelenoMet Nutrient mix from Athena Enzyme Systems) and glucose (5 g l^{-1}) and grown for a further 2 h. IPTG (0.5 m*M*) and a mixture of selenomethionine and methionine (5:1) was added and cultivation was continued for 20 h.

2.2. Protein purification and His-tag cleavage

The cells were harvested by centrifugation and lysed by sonication. The cell debris was removed by centrifugation and the recombinant protein with a six-histidine tag was purified from the lysate using affinity chromatography on an Ni–NTA agarose column (Qiagen). The recombinant protein was eluted with a high imidazole concentration followed by buffer exchange into 20 m*M* Tris–HCl pH 8.0 using an Amicon centrifugal filter unit (Millipore).

The hexahistidine tag was removed from BBA64 by addition of recombinant TEV protease to the protein (5 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0) and incubation for 12 h at room temperature. After cleavage, the protease, digested His tag and remaining uncleaved protein were removed by passage

through an Ni–NTA column (Qiagen). The BBA64 protein was further purified by ion-exchange chromatography on a Mono Q 5/50 GL column (GE Healthcare) linked to an ÄKTA chromatography system (Amersham Biosciences) at a flow rate of 1 ml min^{-1}. The fractions containing pure protein were buffer-exchanged into 20 m*M* Tris–HCl pH 8.0 and concentrated to a concentration of 8 mg ml^{-1} using an Amicon centrifugal filter unit (Millipore).

2.3. Mass spectrometry

To verify SeMet incorporation, MALDI–TOF mass spectrometry was performed for both SeMet-labelled and native protein. 1 µl protein (4 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0) was mixed with 1 µl 0.1% TFA and 1 µl matrix solution consisting of 15 mg ml^{-1} 2,5-dihydroxyacetophenone in 20 m*M* ammonium citrate, 75% ethanol. 1 µl of the obtained mixture was loaded onto the target plate, dried and analyzed using a Bruker Daltonics Autoflex mass spectrometer. The results indicated that the SeMet-labelled protein had a molecular weight that was about 200 Da larger than that of the native protein, which is consistent with the incorporation of all four possible SeMet residues.

The protein state in the crystals was tested using MALDI–TOF mass spectrometry essentially as described above. The obtained protein crystals were dissolved in 20 m*M* Tris–HCl pH 8.0 and compared with the protein batch used for crystallization.

2.4. Estimation of the multimeric state by gel-filtration chromatography

Purified protein sample (5 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0, 0.5 *M* NaCl) was loaded into a pre-packed Superdex 200 10/300 GL column connected to an ÄKTA chromatography system (Amersham Biosciences). The column was pre-equilibrated with the same buffer and run at a flow rate of 0.7 ml min^{-1}. Bovine serum albumin (67 kDa), ovalbumin (43 kDa) and chymotrypsinogen A (25 kDa) were used as molecular-weight reference standards.

2.5. Crystallization of native and SeMet-labelled protein

Crystallization was performed by the sitting-drop vapour-diffusion technique by mixing 1 µl protein solution with an equal volume of precipitant solution. Initial screening was performed using several 96-reagent sparse-matrix screens and a prospective crystal hit was obtained from Structure Screen I + II from Molecular Dimensions. Further optimization of the favourable condition was performed by examining about 300 variant conditions. Needle-shaped crystals with approximate dimensions of 800 × 60 × 20 µm were obtained using a precipitant solution consisting of 20% PEG 2000 MME, 0.05 *M* ammonium sulfate, 15% glycerol. SeMet-labelled protein crystals were obtained under the same conditions. Because of the relatively high concentration of PEG 2000 MME and glycerol in the crystallization mixture, no additional cryoprotection was used before the crystals were flash-cooled in liquid nitrogen.

2.6. Data collection and structure determination

The crystal structure of BBA64 was determined using the multiple-wavelength anomalous diffraction (MAD) technique. Diffraction data for the native and SeMet-labelled proteins were collected on beamline I911-3 at Max-lab (Lund University, Sweden) using a MAR Mosaic 225 detector (MAR Research GmbH). Native crystals of BBA64 diffracted to 2.4 Å resolution and SeMet-labelled crystals diffracted to 2.8 Å resolution. The space group was $P2_12_12_1$ in both cases, with nearly identical unit-cell parameters.

Reflections were indexed and scaled using the *MOSFLM* and *SCALA* programs from the *CCP*4 program suite (Battye *et al.*, 2011; Evans, 2006; Winn *et al.*, 2011). Data were merged using *CAD* (Dodson *et al.*, 1997) and Se-atom positions and initial phases were determined using *SHELXC*/*D*/*E* (Sheldrick, 2008). The initial protein model was built automatically in *Buccaneer* (Cowtan, 2006) and minor rebuilding of the model was performed manually in *Coot* (Emsley & Cowtan, 2004). Water molecules were picked automatically in *Coot* and inspected manually. Crystallographic refinement was carried out with *REFMAC*5 (Murshudov *et al.*, 2011).

3. Results and discussion

3.1. Quality of the model and overall structure

The asymmetric unit contained two protein molecules designated chains *A* and *B*. The protein model of chain *A* was built for residues 92–302 and that of chain *B* for residues 102–297. The signal sequence (residues 1–32) had already been excluded at the expression stage and residues 33–91 could not be seen in the electron-density map. In order to determine whether residues 33–91 were disordered or were absent in the crystal, we performed mass spectrometry of the crystallized material. The results revealed that although present in the purified protein, residues 33–49 were absent in the crystallized material, probably owing to a proteolytic susceptibility as reported for CspA (Cordes *et al.*, 2005). Meanwhile, residues 50–91 could not be seen in the electron-density map probably owing to the flexible nature of the N-terminal part of the protein molecule, which might form an unstructured region as indicated by the secondary-structure prediction software *Jpred* 3 (Cole *et al.*, 2008) and which is likely to serve as a linker between the structured region of the protein and the cell surface.

Residues Asp92–Lys101, Asp216–Pro224 and Leu298–Gln302 in chain *B* were not included in the model owing to weak electron density, although the same regions were well defined in chain *A* and thus no information on protein structure was lost.

The crystal structure of BBA64 reveals the same fold as originally detected in CspA (Cordes *et al.*, 2005) and consists of seven α-helices (from ten to 28 residues in length and named *A*–*G*) crossing at different angles, which are connected by loops of different lengths (Fig. 1*a*). The model also includes five sulfate ions and 114 water molecules. A summary of the

data-collection, refinement and validation statistics is given in Table 1.

3.2. Oligomerization state and C-terminal α-helix

The crystal structure of BBA64 revealed essentially the same overall protein fold as that of CspA, except for the C-terminal part, which has different conformations in the two homologous proteins. In CspA the C-terminal α-helix (α-helix *E*) forms a stalk-like extension that protrudes outwards from the protein molecule and serves as a basis for dimer formation. The C-terminal α-helix *F* in BBA64, which corresponds to α-helix *E* in CspA, instead forms a loop that is not present in CspA and switches backwards, becoming part of a compact α-helical domain. Therefore, the crystal structure of BBA64 clearly indicates that the protein does not form a stable homodimer by interaction of the C-terminal α-helices as in the case of CspA. Since dimerization appears to be essential for the proper function of CspA, as the proposed ligand-binding site was located in the cleft between the two monomers, we performed an additional analysis of the oligomerization state of BBA64. To exclude loss of oligomerization owing to proteolysis, we first confirmed the molecular mass of the BBA64 monomer (peak at 28.7 kDa) by mass spectrometry. We then applied the purified BBA64 protein onto a calibrated gel-filtration column and observed that the protein eluted at between 43 and 25 kDa, indicating that BBA64 (31.6 kDa) is most likely to be a monomer and does not form a stable dimer in a solution.

Furthermore, the interaction of the monomers in the CspA protein leads to the formation of a homodimer in a crystal containing two molecules per asymmetric unit which interact

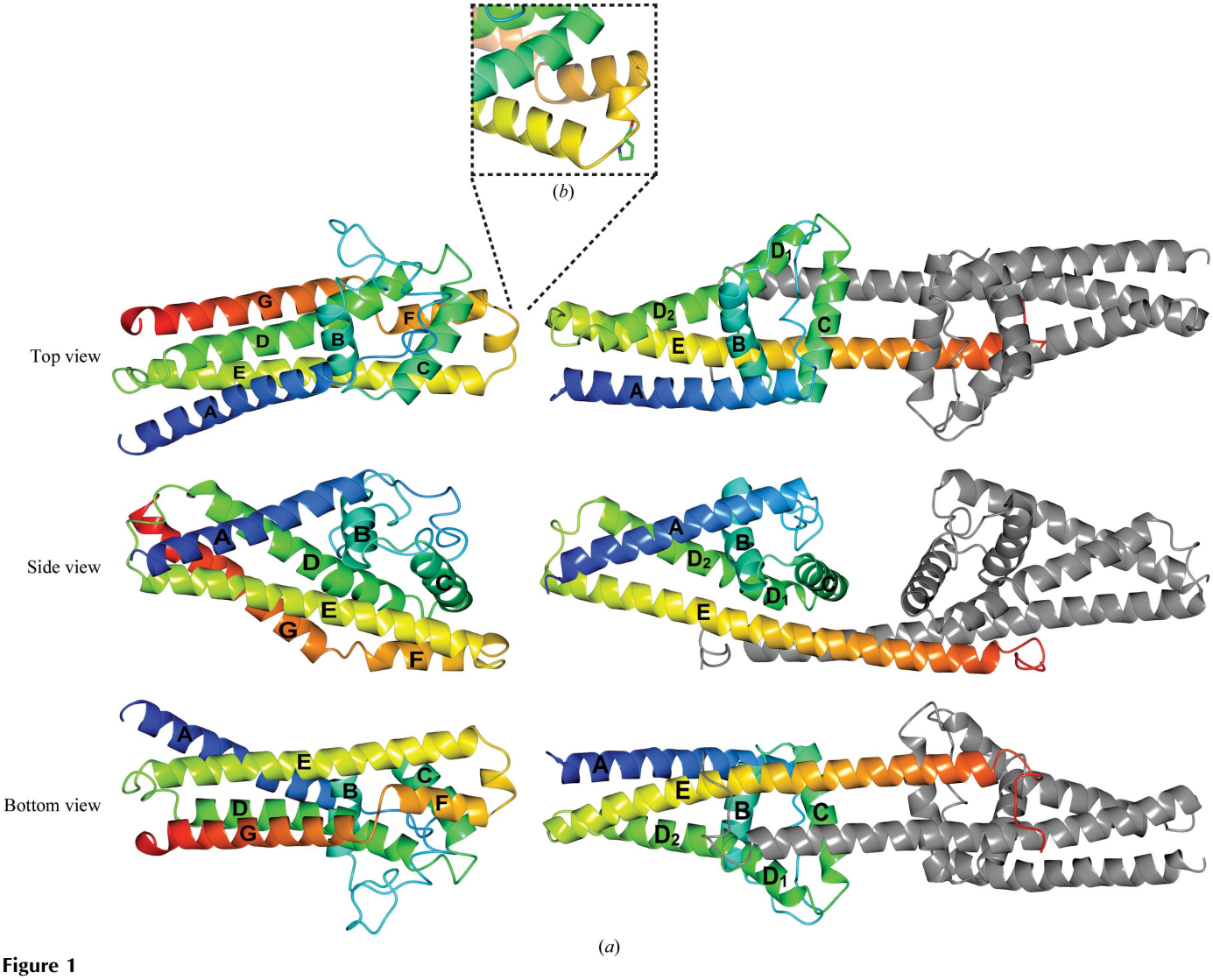

Figure 1
Crystal structures of BBA64 and the orthologous protein CspA. (*a*) A cartoon representation of three different views of BBA64 (left column) and CspA (right column) rotated by 90° in a horizontal plane and coloured using a rainbow colour scheme starting from a blue colour at the N-terminal part and gradually changing to a red colour approaching the C-terminal part of the model. The seven α-helices that form BBA64 are labelled from *A* to *G* starting from the N-terminal part; in CspA the α-helices are named from *A* to *E* starting from the N-terminal part. CspA is represented as a homodimer by overlapping the C-terminal α-helices, forming an extensive 2240.9 Å^2 contact area. (*b*) A loop between α-helices *E* and *F* in BBA64 contains a proline residue and thus limits the flexibility of the loop region.

with their C-terminal α-helices and bury an extensive surface area of 2240.9 Å^2 in the interface site (Fig. 1*b*). This clearly demonstrates that dimer formation is highly probable for CspA. In contrast, although the BBA64 crystal structure also contains two protein molecules in the asymmetric unit, the interface between the two monomers buries a surface area of only 534 Å^2; other possible interfaces in the crystal as defined by the protein interface prediction software *PISA* (Krissinel & Henrick, 2007) bury surface areas of 188–370 Å^2, indicating that stable dimer formation is unlikely for BBA64.

It could be speculated that structural rearrangements in the C-terminal part of BBA64 could occur under certain conditions (*e.g.* at different pH values or temperatures) by switching the folded-back helix into an extended conformation and therefore promoting dimer formation. However, we consider this to be unlikely since a proline residue (Pro258) is located at the end of α-helix *E*, promoting the formation of a turn and limiting the flexibility of the respective helices in relation to each other (Fig. 2*c*).

Table 1
Data and structure-quality statistics.

Values in parentheses are for the highest resolution bin.

		SeMet		
Data set	Native	Peak	Remote	Inflection
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters				
a (Å)	49.15	49.38	49.35	49.11
b (Å)	70.14	69.97	70.09	69.91
c (Å)	186.19	186.50	186.70	186.12
Wavelength (Å)	1.0000	0.9791	0.9686	0.9793
Resolution (Å)	40.00–2.40 (2.46–2.53)	40–2.77 (2.94–2.77)	40–2.74 (2.91–2.74)	40–3.04 (3.23–3.04)
No. of reflections	92002	278029	287525	75123
No. of unique reflections	25348	19973	20587	14969
Completeness (%)	98.2 (99.4)	99.6 (100.0)	99.5 (100.0)	99.2 (100.0)
$R_{\rm merge}$†	0.10 (0.39)	0.18 (0.76)	0.19 (0.83)	0.13 (0.38)
$\langle I/\sigma(I)\rangle$	8.70 (3.1)	12.5 (4.1)	11.8 (3.6)	8.9 (3.7)
Average multiplicity	3.6 (3.6)	14.0 (14.4)	14.0 (14.5)	5.0 (5.2)
Refinement				
$R_{\rm work}$	0.215 (0.255)			
$R_{\rm free}$	0.261 (0.355)			
No. of reflections	24057			
Average *B* factor (Å^2)				
Overall	28.5			
From Wilson plot	33.7			
No. of atoms				
Protein	3231			
Ligand	25			
Water	114			
R.m.s. deviations‡				
Bond lengths (Å)	0.018			
Bond angles (°)	2.060			
Protein geometry§				
Bad rotamers (%)	5.6			
Residues with bad bonds (%)	0.0			
Residues with bad angles (%)	0.0			
Clashscore, all atoms	5.2			
Ramachandran statistics				
Residues in most favoured regions (%)	96.4			
Residues in allowed regions (%)	3.6			
Outliers (%)	0.0			

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity. ‡ R.m.s. deviations from ideal values (Engh & Huber, 1991). § Calculated using *MolProbity* (Chen *et al.*, 2010).

3.3. Sequence comparison

Although dimer formation in CspA may be essential for the proper function of the protein, as has been proposed previously (Cordes *et al.*, 2005, 2006; Kraiczy *et al.*, 2009), there appears to be a contradiction as it has been shown that CspA is able to bind complement factor H and also FHL-1 under denatured conditions after Tris–Tricine SDS–PAGE separation (Kraiczy *et al.*, 2004) and formation of the homodimer has only been observed at very high protein concentrations (Cordes *et al.*, 2005). Therefore, the different orientation of the C-terminal α-helix of BBA64 and its inability to form a dimer may not be the only explanation for the diverse functions of the proteins. Analysis of BBA69, which is another member of the homologous Pfam54 protein family and has the highest sequence identity to CspA among the family members (61% identity), shows that it is thought to form homodimers in the same way as CspA but does not bind either complement factor H or FHL-1. The C-terminal α-helix amino-acid sequence of BBA69 was even made identical to that of CspA using site-directed mutagenesis and indicated that the C-terminal part of CspA is not sufficient to provide binding of complement regulators and that different parts of the protein cooperate to ensure the proper function of CspA (Kraiczy *et al.*, 2009; Wywial *et al.*, 2009). C-terminal deletion mutants of CspA also produced protein which did not bind complement regulators, suggesting that the C-terminal part of the protein could be directly involved in complement factor H and FHL-1 binding or could be responsible for the stability of the overall protein fold, as the functional assays for CspA were performed under reduced conditions (Wallich *et al.*, 2005). Therefore, to obtain further understanding about the differences in the structures and functions of the respective proteins, we performed a structure-based sequence alignment to highlight their differences and to indicate their importance in relation to the different appearances of the proteins (Fig. 2*a*).

Three potential sites on CspA which are essential for the function of the protein have been predicted (Wallich *et al.*, 2005; Kraiczy *et al.*, 2004, 2009; McDowell *et al.*, 2005). One of the sites, which is proposed to be involved in binding complement factor H and factor-H-like protein-1 (FHL-1), resides on α-helix *C* (residues 145–154) located at the cleft between the two molecules if a homodimer is formed. The

second and third sites are located at the C-terminal α-helix *E* (residues 204–213 and 233–242) which constitutes the required surface area for dimer formation; interruption of this region could disrupt the formation of the dimers or affect the correct folding of a molecule and thus affect the function of the protein (Fig. 2*d*). By using mutational analysis, a number of specific residues located at these sites on CspA helices *C* and *E* have been identified which are relevant for dimer formation or interaction with FHL-1 and factor H binding proteins (Kraiczy *et al.*, 2009). Therefore, we searched for the respective residue locations in BBA64 using structure-based sequence alignment in order to justify the differences in protein function.

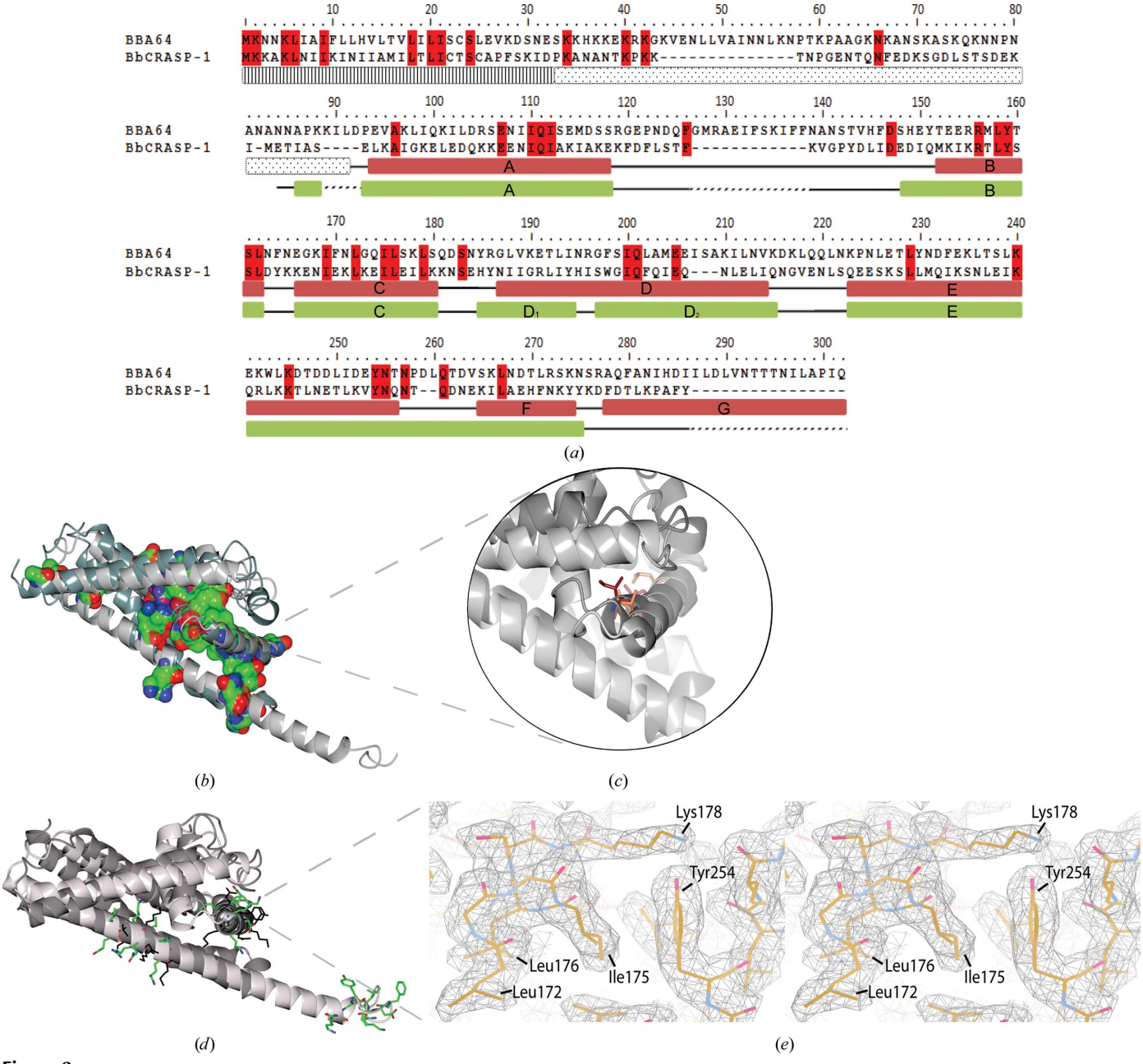

Figure 2
Structure-based sequence alignment of the homologous proteins BBA64 and CspA. (*a*) The initial alignment was obtained using the *PRALINE* multiple sequence-alignment tool and was adjusted manually based on the three-dimensional structures. Conserved residues are highlighted in red. The signal sequence region for BBA64 is shown as a rectangle filled with vertical stripes and the N-terminal part of BBA64 which cannot be seen in the electron-density map is shown as a dotted rectangle. Secondary-structure elements as deduced from the crystal structures of the BBA64 and CspA proteins are represented as rectangles for α-helices and as lines for loop regions and are shown below the sequence alignment. (*b*) Location of the conserved residues in superimposed crystal structures of the two homologous proteins BBA64 (dark grey) and CspA (light grey). Conserved amino acids are shown using a surface-representation model of the residues. (*c*) Conserved hydrophobic residues are directed inwards towards the protein core and are not exposed on the surface, suggesting that they are necessary for correct folding of the molecule and are relevant for α-helix orientation. (*d*) Superimposed crystal structures of BBA64 (dark grey with residues in black) and CspA (light grey with residues in green) showing the residues in three regions identified in CspA as being necessary for correct binding of complement factor H and FHL-1. (*e*) Stereoview of the $2mF_o - DF_c$ electron-density map contoured at 1.0σ located on α-helix *C* of BBA64 and representing protein residues as stick models.

In the first region, which is thought to be involved in the binding of complement regulators, residues Lys136–Glu147 in α-helix *C* which forms the cleft in the dimer were systematically mutated, leading to only a slightly lower complement binding except for the residue Leu146, the mutation of which to a histidine led to a significant decrease in the binding of both complement factor H and FHL-1. Comparison of the amino-acid sequence of the potential binding site located at α-helix *C* of CspA with the same region of BBA64 revealed that Leu146 is conserved in the two homologues, together with other hydrophobic residues including Ile139, Leu142, Ile145 and Leu149. Structural analysis of the respective region of BBA64 and CspA reveals that all five conserved amino acids are located in the hydrophobic core of the protein monomer at the interface of α-helices *C* and *D*, and are not exposed at the surface of a monomer (Fig. 2*c*). Considering that all of these residues are hydrophobic, it can be assumed that they are necessary for the stability of the core and that the mutation of any of these residues could dislocate α-helix *C* or simply lead to incorrectly folded nonfunctional protein. Therefore, the difference in the nonconserved residues located on the surface of the two protein molecules, in this case on α-helix *C*, could explain the divergent characters and functions of the homologous proteins.

The second region of CspA, which is involved in the formation of the dimer interface located in the C-terminal part of the protein, was also exposed to mutational analysis. The crystal structure of BBA64 reveals a different conformation to that of CspA in that the C-terminal α-helix does not form a stalk-like extension running outwards from the central part of the protein and thus does not support dimer formation by itself, and the residues that were detected to be critical for dimer formation (Tyr240, Asp242 and Leu246) at the end of the C-terminal α-helix of CspA are substituted by other residues in BBA64, which is consistent with its monomeric state. In the C-terminal α-helix of BBA64, in addition to the

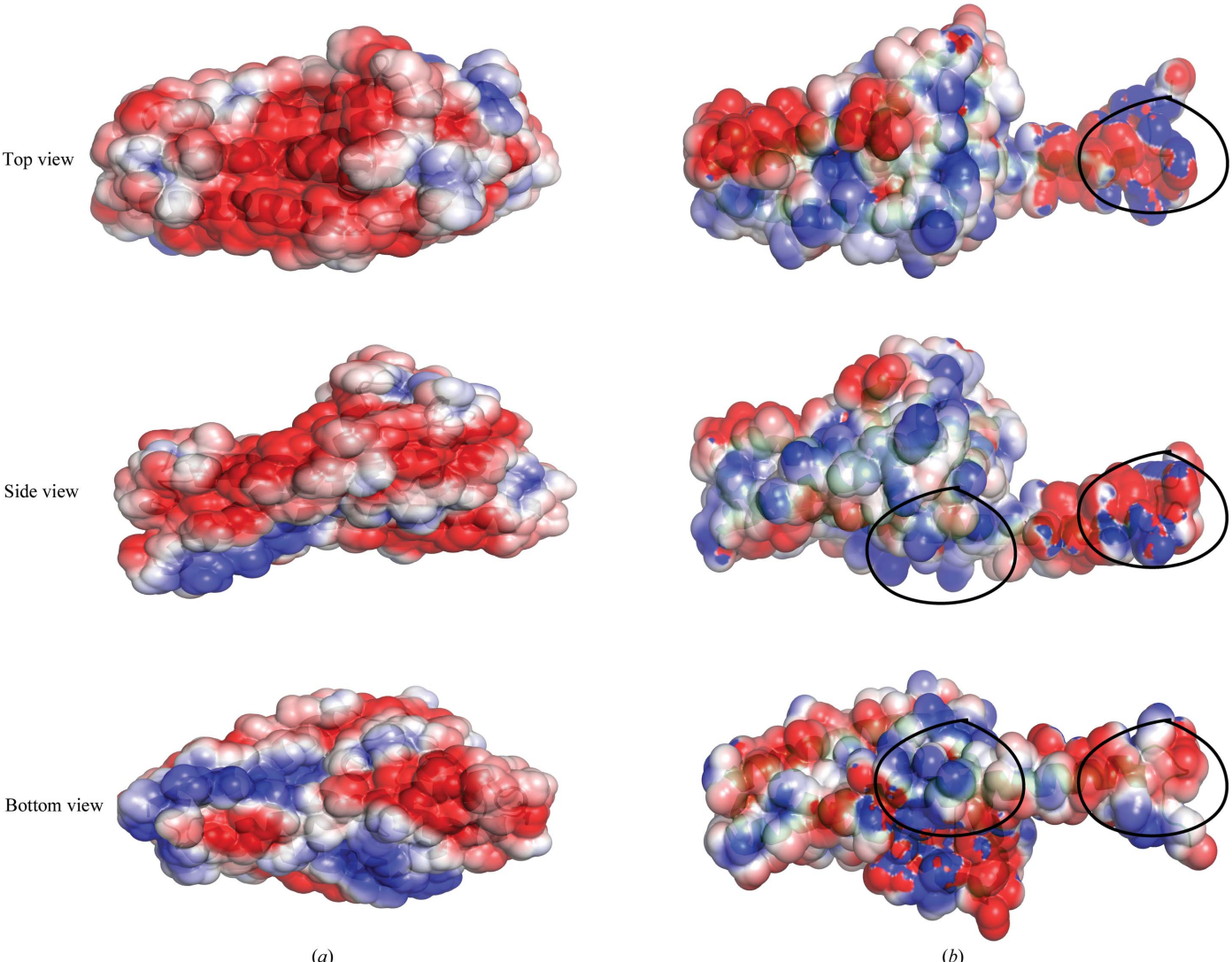

Figure 3
Electrostatic potential surface of BBA64 and CspA. The BBA64 monomer (*a*) is shown in the same orientation and on the same scale as the CspA monomer (*b*). Top, side and bottom views are shown in the same orientation as in Fig. 1. The electrostatic potential (red, negative; blue, positive) was calculated using *APBS* (Baker *et al.*, 2001).

substitution of the residues relevant to dimer formation, there is an extension of 16 residues in the C-terminal α-helix *G*. The importance of this extension can be simply understood from the crystal structure: the α-helix interacts with the other α-helices *D* and *F* in BBA64 covering the whole length and forming a compact structure; in addition, it buries several hydrophobic residues at the interface between the helices, forming a more stable folded-back conformation. By analyzing the three different parts of CspA important for binding factor H and FHL-1 and comparing them with the same regions in BBA64, it becomes clear that although CspA can bind complement regulators as a monomer (as judged from its ability to bind complement factor H and FHL-1 after SDS–PAGE separation), the three regions that are necessary for proper binding exhibit considerable sequence diversity between the two homologous proteins and thus could also explain the inability of BBA64 to bind complement regulators in addition to the different fold of the C-terminal α-helix.

Interestingly, one general difference that can be noticed on analyzing the sequences and superimposed crystal structures is located in the loop region between α-helices *A* and *B* of BBA64 and CspA (Fig. 1). The loop in CspA consists of 17 residues and is remarkably shorter than that in BBA64, where the loop is formed of 29 residues, making this region of the protein more flexible. In addition, structure analysis using the *PISA* (*Protein Interfaces, Surfaces and Assemblies*) prediction tool (Krissinel & Henrick, 2007) reveals that the largest possible interface region in BBA64 that could interact with other molecules buries 534 Å^2 of surface area and also involves a large portion of the loop region between α-helices *A* and *B*. Therefore, we speculate that the loop region in BBA64 is likely to be related to the potential protein function and could be involved in formation of the protein binding site.

3.4. Overall electrostatic properties of BBA64 and CspA

In addition to comparison of the overall protein fold and amino-acid sequence, we also performed a comparison of the electrostatic properties of the two homologous proteins. The charge distribution of a molecule can influence the binding of a ligand or receptor, especially if the contact area between the two molecules buries a large surface area, as could be the case in CspA if we assume that all three of the regions described above are indeed necessary for binding complement factor H and FHL-1. From the electrostatic potential surface of BBA64 it can be deduced that the dominant charge of the molecule is negative and that only a small region at the bottom of the molecule has a positive surface potential (Fig. 3). The surface potentials of BBA64 and CspA could also explain the difference in their respective binding partners, as the surface potential of the segments in the CspA molecule which are thought to be necessary for the binding of complement factor H and FHL-1 (marked by circles in Fig. 3) are mainly positively charged, which is the opposite of the same segments in BBA64.

3.5. Potential function of BBA64

With regard to the potential ligand or receptor for BBA64, it has been shown that after treatment of *B. burgdorferi* with anti-BBA64 (an antibody for the outer surface protein BBA64) the binding of the bacteria to HUVEC (human umbilical vein endothelial cells) or H4 (human neuroglial cells) cells is significantly reduced, indicating that the potential binding partner that provides adherence and invasion of the Lyme disease agent could be located on the cell surface of the mammalian host organism (Schmit *et al.*, 2011). Considering that lipoprotein BBA64 is necessary to provide the transfer of *Borrelia* from tick salivary glands to the host organism, the binding of *Borrelia* to mammalian cells described above could reduce the range of the possible ligand/receptor search, leading to the exclusion of, for example, tick salivary gland proteins, which could otherwise be potential ligand candidates for consideration in assisting in the transport of *B. burgdorferi* to the host organism. The location of the receptors on the surface of the host cells could also explain the vital role of the protein in the transfer of *B. burgdorferi* from ticks to mammals and the observation that *Borrelia* lacking BBA64 synthesis was unable to infect mice *via* tick feeding but that infection could be initiated after needle inoculation of *B. burgdorferi* (Gilmore *et al.*, 2010; Patton *et al.*, 2011). It should be noted that the expression of the outer surface lipoprotein BBA64 is also highly upregulated in tissues during chronic murine infection throughout the infection period, indicating that the protein not only contributes to the transfer of *Borrelia* from ticks to the host organism and the establishment of the infection but also to the maintenance and persistence of Lyme disease (Gilmore *et al.*, 2008).

4. Conclusions

The crystal structure of *B. burgdorferi* outer surface lipoprotein BBA64, which plays a vital role in the pathogenesis of Lyme disease, has been solved at 2.4 Å resolution. There are several differences between BBA64 and the related homologous CspA protein, such as surface-exposed residue divergence and dissimilar electrostatic surface potentials, that could explain the distinct functions of these proteins. However, the main difference is the orientation of the C-terminal α-helix, which in BBA64 does not form a stalk-like extension running outwards from the protein and does not promote the formation of a homodimer, while in CspA the C-terminal part has been shown to be essential for the correct function of the protein and is involved in dimer formation. Although the exact function of BBA64 is still unclear, taking into account that the protein is a relevant component of *B. burgdorferi* pathogenesis the determination of the crystal structure is an important step forward to help to clarify the exact binding partner of the protein and to contribute to drug development against Lyme disease.

Coordinates and structure factors have been deposited in the Protein Data Bank (Berman *et al.*, 2000) with accession code 4aly.

This work was supported by ESF grant 1DP/1.1.1.2.0/09/APIA/VIAA/150 and Latvian Council of Science grant No. 10.0029.3. We thank Dr Gunter Stier and Dr Huseyin Besir from EMBL for providing the expression vector pETm-11. We also acknowledge the staff at the MAX-lab synchrotron for their support during data collection.

References

Angel, T. E., Luft, B. J., Yang, X., Nicora, C. D., Camp, D. G. II, Jacobs, J. M. & Smith, R. D. (2010). *PLoS One*, **5**, e13800.

Anguita, J., Samanta, S., Revilla, B., Suk, K., Das, S., Barthold, S. W. & Fikrig, E. (2000). *Infect. Immun.* **68**, 1222–1230.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.

Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. (2004). *J. Mol. Biol.* **340**, 783–795.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Brooks, C. S., Hefty, P. S., Jolliff, S. E. & Akins, D. R. (2003). *Infect. Immun.* **71**, 3371–3383.

Burgdorfer, W., Barbour, A. G., Hayes, S. F., Benach, J. L., Grunwaldt, E. & Davis, J. P. (1982). *Science*, **216**, 1317–1319.

Carroll, J. A., Cordova, R. M. & Garon, C. F. (2000). *Infect. Immun.* **68**, 6677–6684.

Casjens, S. R. *et al.* (2012). *PLoS One*, **7**, e33280.

Casjens, S., Palmer, N., van Vugt, R., Huang, W. M., Stevenson, B., Rosa, P., Lathigra, R., Sutton, G., Peterson, J., Dodson, R. J., Haft, D., Hickey, E., Gwinn, M., White, O. & Fraser, C. M. (2000). *Mol. Microbiol.* **35**, 490–516.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Cole, C., Barber, J. D. & Barton, G. J. (2008). *Nucleic Acids Res.* **36**, W197–W201.

Cordes, F. S., Kraiczy, P., Roversi, P., Simon, M. M., Brade, V., Jahraus, O., Wallis, R., Goodstadt, L., Ponting, C. P., Skerka, C., Zipfel, P. F., Wallich, R. & Lea, S. M. (2006). *Int. J. Med. Microbiol.* **296**, Suppl. 40, 177–184.

Cordes, F. S., Kraiczy, P., Roversi, P., Skerka, C., Kirschfink, M., Simon, M. M., Brade, V., Lowe, E. D., Zipfel, P., Wallich, R. & Lea, S. M. (2004). *Acta Cryst.* D**60**, 929–932.

Cordes, F. S., Roversi, P., Kraiczy, P., Simon, M. M., Brade, V., Jahraus, O., Wallis, R., Skerka, C., Zipfel, P. F., Wallich, R. & Lea, S. M. (2005). *Nature Struct. Mol. Biol.* **12**, 276–277.

Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.

Dodson, E. J., Winn, M. & Ralph, A. (1997). *Methods Enzymol.* **277**, 620–633.

Embers, M. E., Ramamoorthy, R. & Philipp, M. T. (2004). *Microbes Infect.* **6**, 312–318.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Fraser, C. M. *et al.* (1997). *Nature (London)*, **390**, 580–586.

Gilmore, R. D., Howison, R. R., Dietrich, G., Patton, T. G., Clifton, D. R. & Carroll, J. A. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 7515–7520.

Gilmore, R. D., Howison, R. R., Schmit, V. L. & Carroll, J. A. (2008). *Microb. Pathog.* **45**, 355–360.

Gilmore, R. D., Howison, R. R., Schmit, V. L., Nowalk, A. J., Clifton, D. R., Nolder, C., Hughes, J. L. & Carroll, J. A. (2007). *Infect. Immun.* **75**, 2753–2764.

Hughes, J. L., Nolder, C. L., Nowalk, A. J., Clifton, D. R., Howison, R. R., Schmit, V. L., Gilmore, R. D. & Carroll, J. A. (2008). *Infect. Immun.* **76**, 2498–2511.

Indest, K. J., Ramamoorthy, R., Solé, M., Gilmore, R. D., Johnson, B. J. & Philipp, M. T. (1997). *Infect. Immun.* **65**, 1165–1171.

Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H. & Krogh, A. (2003). *Protein Sci.* **12**, 1652–1662.

Kraiczy, P., Hanssen-Hübner, C., Kitiratschky, V., Brenner, C., Besier, S., Brade, V., Simon, M. M., Skerka, C., Roversi, P., Lea, S. M., Stevenson, B., Wallich, R. & Zipfel, P. F. (2009). *Int. J. Med. Microbiol.* **299**, 255–268.

Kraiczy, P., Hellwage, J., Skerka, C., Becker, H., Kirschfink, M., Simon, M. M., Brade, V., Zipfel, P. F. & Wallich, R. (2004). *J. Biol. Chem.* **279**, 2421–2429.

Kraiczy, P., Rossmann, E., Brade, V., Simon, M. M., Skerka, C., Zipfel, P. F. & Wallich, R. (2006). *Wien. Klin. Wochenschr.* **118**, 669–676.

Kraiczy, P., Skerka, C., Kirschfink, M., Brade, V. & Zipfel, P. F. (2001). *Eur. J. Immunol.* **31**, 1674–1684.

Kraiczy, P., Skerka, C., Kirschfink, M., Zipfel, P. F. & Brade, V. (2001). *Int. Immunopharmacol.* **1**, 393–401.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

McDowell, J. V., Harlin, M. E., Rogers, E. A. & Marconi, R. T. (2005). *J. Bacteriol.* **187**, 1317–1323.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Ojaimi, C., Brooks, C., Casjens, S., Rosa, P., Elias, A., Barbour, A., Jasinskas, A., Benach, J., Katona, L., Radolf, J., Caimano, M., Skare, J., Swingle, K., Akins, D. & Schwartz, I. (2003). *Infect. Immun.* **71**, 1689–1705.

Patton, T. G., Dietrich, G., Dolan, M. C., Piesman, J., Carroll, J. A. & Gilmore, R. D. (2011). *PLoS One*, **6**, e19536.

Ramamoorthy, R. & Scholl-Meeker, D. (2001). *Infect. Immun.* **69**, 2739–2742.

Revel, A. T., Talaat, A. M. & Norgard, M. V. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 1562–1567.

Schmit, V. L., Patton, T. G. & Gilmore, R. D. (2011). *Front. Microbiol.* **2**, 141.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Steere, A. C., Coburn, J. & Glickstein, L. (2004). *J. Clin. Invest.* **113**, 1093–1101.

Tokarz, R., Anderton, J. M., Katona, L. I. & Benach, J. L. (2004). *Infect. Immun.* **72**, 5419–5432.

Wallich, R., Pattathu, J., Kitiratschky, V., Brenner, C., Zipfel, P. F., Brade, V., Simon, M. M. & Kraiczy, P. (2005). *Infect. Immun.* **73**, 2351–2359.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Wywial, E., Haven, J., Casjens, S. R., Hernandez, Y. A., Singh, S., Mongodin, E. F., Fraser-Liggett, C. M., Luft, B. J., Schutzer, S. E. & Qiu, W.-G. (2009). *Gene*, **445**, 26–37.