the FEBS Journal

Structural characterization of CspZ, a complement regulator factor H and FHL-1 binding protein from *Borrelia burgdorferi*

Kalvis Brangulis[1,2], Ivars Petrovskis[1], Andris Kazaks[1], Janis Bogans[1], Martins Otikovs[4], Kristaps Jaudzems[4], Renate Ranka[1,2] and Kaspars Tars[1,3]

1 Latvian Biomedical Research and Study Centre, Riga, Latvia
2 Riga Stradins University, Latvia
3 University of Latvia, Riga, Latvia
4 Latvian Institute of Organic Synthesis, Riga, Latvia

Keywords
complement system; immune system; *Ixodes* ticks; Lyme borreliosis; outer surface proteins

Correspondence
K. Brangulis, Latvian Biomedical Research and Study Centre, Ratsupites Street 1, LV-1067, Riga, Latvia
Fax: +371 6744 2407
Tel: +371 6780 8200
E-mail: kalvis.brangulis@inbox.lv

(Received 24 October 2013, revised 29 March 2014, accepted 4 April 2014)

doi:10.1111/febs.12808

Borrelia burgdorferi is the causative agent of Lyme disease and is found in two different types of hosts in nature – *Ixodes* ticks and various mammalian organisms. To initiate disease and survive in mammalian host organisms, *B. burgdorferi* must be able to transfer to a new host, proliferate, attach to different tissue and resist the immune response. To resist the host's immune response, *B. burgdorferi* produces at least five different outer surface proteins that can bind complement regulator factor H (CFH) and/or factor H-like protein 1 (CFHL-1). The crystal structures of two uniquely folded complement binding proteins, which belong to two distinct gene families and are not found in other bacteria, have been previously described. The crystal structure of the CFH and CFHL-1 binding protein CspZ (also known as BbCRASP-2 or BBH06) from *B. burgdorferi*, which belongs to a third gene family, is reported in this study. The structure reveals that the overall fold is different from the known structures of the other complement binding proteins in *B. burgdorferi* or other bacteria; this structure does not resemble the fold of any known protein deposited in the Protein Data Bank. The N-terminal part of the CspZ protein forms a four-helix bundle and has features similar to the FAT domain (focal adhesion targeting domain) and a related domain found in the vinculin/α-catenin family. By combining our findings from the crystal structure of CspZ with previous mutagenesis studies, we have identified a likely binding surface on CspZ for CFH and CFHL-1.

Introduction

Lyme disease is a multisystemic infection that is primarily caused by five spirochete genospecies that belong to the *Borrelia burgdorferi sensu lato* complex: *B. burgdorferi sensu stricto*, *B. garinii*, *B. afzelii*, *B. spielmanii* and *B. bavariensis* [1–3]. *Ixodes* ticks act as vectors to transmit Lyme disease causing bacteria between different mammalian hosts. Therefore, *B. burgdorferi* must survive in various conditions found in both the *Ixodes* ticks and mammalian hosts. To induce an infection, the bacteria must be

Abbreviations
BbCRASP, *Borrelia burgdorferi* complement regulator-acquiring surface protein; CFH, complement regulator factor H; CFHL-1, complement factor H-like protein 1; CRASP, complement regulator-acquiring surface protein; CspZ, BbCRASP-2; FAT, focal adhesion targeting; SCR, short consensus repeat; Se-Met, L-selenomethionine.

transferred from infected *Ixodes* ticks to mammals during a blood meal. Upon transfer to the new host organism, *B. burgdorferi* survives by resisting the host's immune response and then proliferates and spreads throughout the body. To resist the host's immune response, *B. burgdorferi* binds to the major alternative pathway complement regulator factor H (CFH) and factor H-like protein 1 (CFHL-1) [4]. These two regulatory proteins are part of the alternative pathway of the complement system. They are related glycoproteins that contain globular domains known as Sushi domains or short consensus repeats (SCRs) [5]. CFH (150 kDa protein) contains 20 SCR domains connected by short loop regions, and CFHL-1 (42 kDa protein) comprises seven SCR domains that are similar to the first seven SCR domains found in CFH [6]. Both proteins inhibit complement activation by accelerating the decay of the C3 convertase complex in the alternative pathway and act as cofactors for factor-I-mediated cleavage of C3b [7].

By binding CFH and CFHL-1, the bacteria mimic the action of the host's own cells, which bind the complement regulators using glycosaminoglycans and heparin which are located on the cell surface, thus avoiding the host's complement attack [8,9].

The proteins that bind to CFH and CFHL-1 in *Borrelia* are termed complement regulator-acquiring surface proteins or CRASPs [10,11]. In *B. burgdorferi*, five CRASP proteins have been identified, and they belong to three different gene families [12]. CRASPs (or BbCRASPs in the case of *B. burgdorferi*) differ in their primary and tertiary protein structures and possess different binding modes for CFH and CFHL-1. CspA (BbCRASP-1) belongs to the PFam54 gene family and can bind to both CFH and CFHL-1 [13]. The crystal structure of CspA was previously determined, and a putative binding cleft in the homodimer has been proposed as a binding site for complement regulators (PDB entry 1W33) [14,15]. CspZ (BbCRASP-2) does not share any sequence similarities with other proteins, including the other CRASPs expressed in *B. burgdorferi*; however, this protein does have the ability to bind both CFH and CFHL-1 [16,17]. ErpP (BbCRASP-3), ErpC (BbCRASP-4) and ErpA (BbCRASP-5) belong to the OspE/F-related (Erp) protein family. These proteins display high sequence similarity with each other and bind only to CFH [18–20]. The crystal structure for ErpC has been determined in complex with CFH domains 19–20 (PDB entry 4J38) [21].

The determination of the CspZ crystal structure reveals a third protein family that is responsible for the binding of CFH and CFHL-1 in *B. burgdorferi*, and this analysis provides more complete insight into the structural biology of the CRASP proteins.

Results and discussion

Crystal structure of CspZ

The final model of CspZ contains one molecule per asymmetric unit. Most of the 214 amino acid residues are accounted for in the protein molecule, including a residual N-terminal glycine residue that remained after the cleavage of the 6xHis purification tag. The electron density map was poor in the loop region between α-helices αB and αC, where four residues were not modeled, and also in the loop between αE and αF, where one residue was not placed.

Overall, the CspZ molecule consists of a single domain, and the molecule appears to have a single hydrophobic core. However, the molecule can also be regarded as being built from two lobes, an N-terminal lobe and a C-terminal lobe, as concluded from the DALI and PDBEFOLD analysis [22,23].

Both lobes consist almost entirely of α-helices. The fold for the N-terminal lobe is similar to a right-turn four-helix bundle. However, except for αE, the α-helices are slightly bent, and there are short loops between helices αA and αB, as well as between αC and αD, which are usually continuous helices in a standard four-helix bundle domain (Fig. 1).

The C-terminal lobe is made from three α-helices. Helices αG and αH in the C-terminal lobe are oriented in an antiparallel fashion and are connected by a short loop region, forming a helix–turn–helix motif. The third helix αI in the C-terminal lobe is connected to αH by a long loop region, resulting in αC running parallel to αH (Fig. 1).

The N-terminal lobe is a potential binding site for CFH and CFHL-1, as identified in previous mutagenesis studies and sequence comparisons with homologous CspZ proteins that do not bind CFH and CFHL-1, which are found in several other *B. burgdorferi* strains [16,17,24,25].

Analysis of the primary sequence by the SIB BLAST network service and INTERPROSCAN [26,27] did not reveal any similarities with non-CspZ proteins, including the other known CFH and/or CFHL-1 binding proteins.

DALI server analysis (three-dimensional alignment with other protein structures) also failed to find any similarities with other CFH and/or CFHL-1 binding proteins, which suggests that the fold of CspZ is unique among the CFH/CFHL-1 binding proteins. However, a search of the DALI server and PDBEFOLD

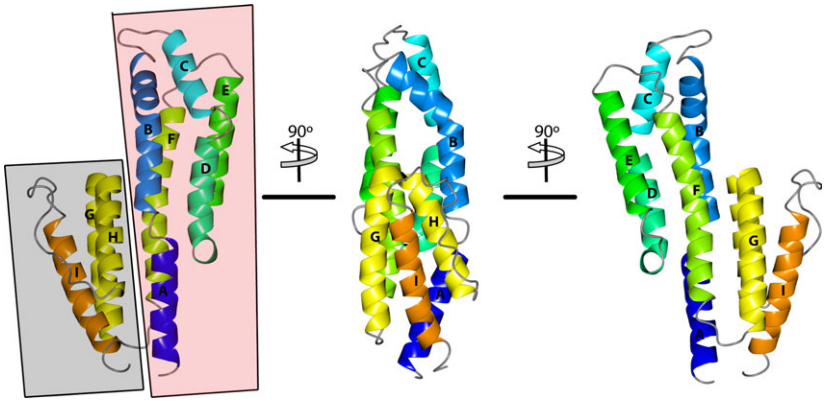

Fig. 1. The overall structure of CspZ. A cartoon representation of three different orientations of CspZ, rotated by 90° in the vertical plane. The N-terminal and C-terminal lobes are shown with pink and grey backgrounds, respectively. The molecules follow a chainbow, starting with a blue color at the N-terminus and gradually changing to red at the C-terminus. The α-helices in CspZ are labeled from A to I, starting from the N-terminus.

[22,23] did find similarities to other proteins that have functions unrelated to the binding of complement factors. These three-dimensional similarities were found only in the N-terminal lobe of the CspZ protein, and the proteins identified by the DALI server are presented in Table 1. The similarities to the mentioned proteins occur within a four-helix bundle recognized as a FAT domain (focal adhesion targeting domain) or a FAT-related domain that is found in the vinculin/α-catenin family.

FAT domains are found in focal adhesion kinases that are involved in integrin-mediated signaling [28]. The FAT domain plays an important role in relaying the signals that are generated by the attachment of the cells to the extracellular matrix. These signals are transmitted through the integrins to cytoplasmic and nuclear targets and regulate the migration, survival and proliferation of the cells [28,29].

Although the three-dimensional structure of the CspZ N-terminal lobe is similar to the aforementioned proteins and the crystal structures can be superimposed, none of these molecules has regions that show any sequence similarity to CspZ. Although the only annotated function for CspZ is the binding of CFH/CFHL-1, it has been shown that CspZ also binds to several approximately 60 kDa serum proteins from different mammalian species, but the identities of these proteins have not been determined [17]. Whether the aforementioned structural similarities are relevant to the actual function of the N-terminal lobe remains to

Table 1. Properties and functions of proteins with similar folds to CspZ N-terminal lobe.

No.	Protein	Function	PDB ID	*Z* score	RMSD (Å)	Identity (%)	*N* residues
1.	Breast cancer anti-estrogen resistance protein 1 from *Homo sapiens*	Serves as a docking protein involved in tyrosine kinase based signaling related to cell adhesion and cell migration [41]	3T6G chain B	9.9	2.6	11	104
2.	Alpha E-catenin from *Mus musculus*	Cell–cell adhesion, links the cadherin–catenin complex to the actin cytoskeleton [42,43]	1L7C	9.9	3.0	9	103
3.	Programmed cell death protein 10 from *H. sapiens*	Required for normal cardiovascular development, involved in multiple signaling pathways, binds to paxillin (signal transduction adaptor protein) [35,36]	3RQE	9.6	2.9	7	96
4.	Vinculin from *Gallus gallus*	Focal adhesion and adherens junctions by binding the actin filaments [33]	1ST6	9.0	2.7	13	94
5.	Talin-1 from *M. musculus*	Adaptor protein that couples the integrin family of cell adhesion molecules to cytoskeletal actin [34]	2X0C	8.1	2.5	3	90

be determined, as other functions, such as binding to different serum proteins, the extracellular matrix or host cells, cannot be excluded. It should be mentioned that the CFH/CFHL-1-binding protein CspA from *B. burgdorferi* is a multifunctional outer surface protein of *B. burgdorferi* and can bind to laminin, plasminogen, fibronectin, collagen and bone morphogenic protein 2 [30]. Although only as an assumption, the indicator that CspZ reveals structural similarities with certain mammalian proteins involved in adhesion and that some of the CspZ orthologs does not bind CFH/CFHL-1, can be an indicator for another multifunctional protein of *B. burgdorferi*.

Residues in CspZ that are potentially involved in CFH/CFHL-1 binding

Several studies of the CspZ protein have investigated the residues that are potentially involved in binding to CFH/CFHL-1 [16,24,25]. With the crystal structure in hand, it is now possible to map the binding site in CspZ and evaluate the potential residues that were previously shown to be involved in CFH/CFHL-1 binding.

Using the PEPSPOT analysis created from the library of peptides representing the CspZ protein, the four regions involved in the binding of CFH and CFHL-1 were previously identified in CspZ by Hartmann *et al.* [16]. The first region spanned residues 34–52, the second region involved residues 70–88, the third region contained residues 127–145, and the fourth region involved residues 202–226. The second region bound only to CFHL-1, while the other regions bound both of the complement factors. Additionally, the last 16 C-terminal residues of CspZ were shown to be important for binding to complement regulators, as the deletion of more than 16 residues from the C-terminus abolished the binding. The crystal structure of CspZ revealed that the residues in these four regions occupy a large surface area on all sides of the molecule, making it difficult to determine the exact binding position of CFH/CFHL-1 in the CspZ protein (Fig. 2A). It is possible that the SCR domains found in CFH/CFHL-1 somehow wrap around the CspZ molecule, forming an extensive contact area between the two molecules. The loss of binding to the complement factors resulting from the truncation of the CspZ C-terminus could be attributed to its subsequent misfolding, as the C-terminal and N-terminal lobes have a common hydrophobic core.

To understand the importance of the regions described by Hartmann *et al.*, an alanine-scanning mutagenesis approach was used by Siegel *et al.*, in which the residues located in each of the four regions were replaced with alanine to determine the role of each specific residue in CFH/CFHL-1 binding [25]. When mutating the residues located in the second region (residues 70–88) (Fig. 2A), the largest effect on CFH/CFHL-1 binding was observed after the mutation of residue Phe81 to alanine. Binding also decreased when residues Gln71, Asn75, Ser79 and Asp84 were replaced with alanine. The loss of binding capacity after replacing Phe81 with alanine might also arise from misfolding, as Phe81 is located near the protein core. Residues Gln71, Asn75, Ser79 and Asp84 form hydrogen bonds with other residues and are involved in the stabilization of the local fold (Fig. 2B).

In the third and fourth regions, several residues that reduced the binding of CFH/CFHL-1 were detected by Siegel *et al.* In the third region (amino acids 127–145), alanine substitution of residues Asp133, Lys136, Lys137 and Lys 140 did not affect binding, and only a slight decrease in CFH/CFHL-1 binding was observed for residues Arg129, Arg139 and His130. Mutations in the fourth region (amino acids 202–226) significantly impaired or completely abolished the binding of CFH/CFHL-1. Inspection of the fourth region in the crystal structure of CspZ revealed that the residues are largely located on the surface of the protein and are presumably not important for the stabilization of the overall fold (Fig. 2B). Notably, hydrophobic residue Phe210 is clearly located on the surface of the CspZ molecule, and mutation of it to alanine resulted in significantly reduced binding to CFH/CFHL-1. The presence of hydrophobic residues on the surface of the protein molecule decreases the conformational stability of the protein because it disrupts the hydration structure that is mediated by hydrogen bonds on the surface of the molecule [31]. Therefore, this residue probably meets a functional need and may be involved in the binding of CFH/CFHL-1. Replacement of the hydrophilic residue Arg204 with alanine resulted in a significant reduction in CFH/CFHL-1 binding; this residue is also located on the surface of the CspZ molecule and could be involved in CFH/CFHL-1 binding. Alanine substitution of residues Arg206, Tyr207 and Glu214 resulted in significantly impaired binding to CFH/CFHL-1, and these residues are most probably not involved in CFH/CFHL-1 binding but rather participate in hydrogen bonding with other nearby residues to stabilize the local fold of the CspZ protein. For example, Tyr207 forms hydrogen bonds with Glu186 and Arg206, and Arg206 forms a hydrogen bond with Cys187 from α-helix H (Fig. 2C).

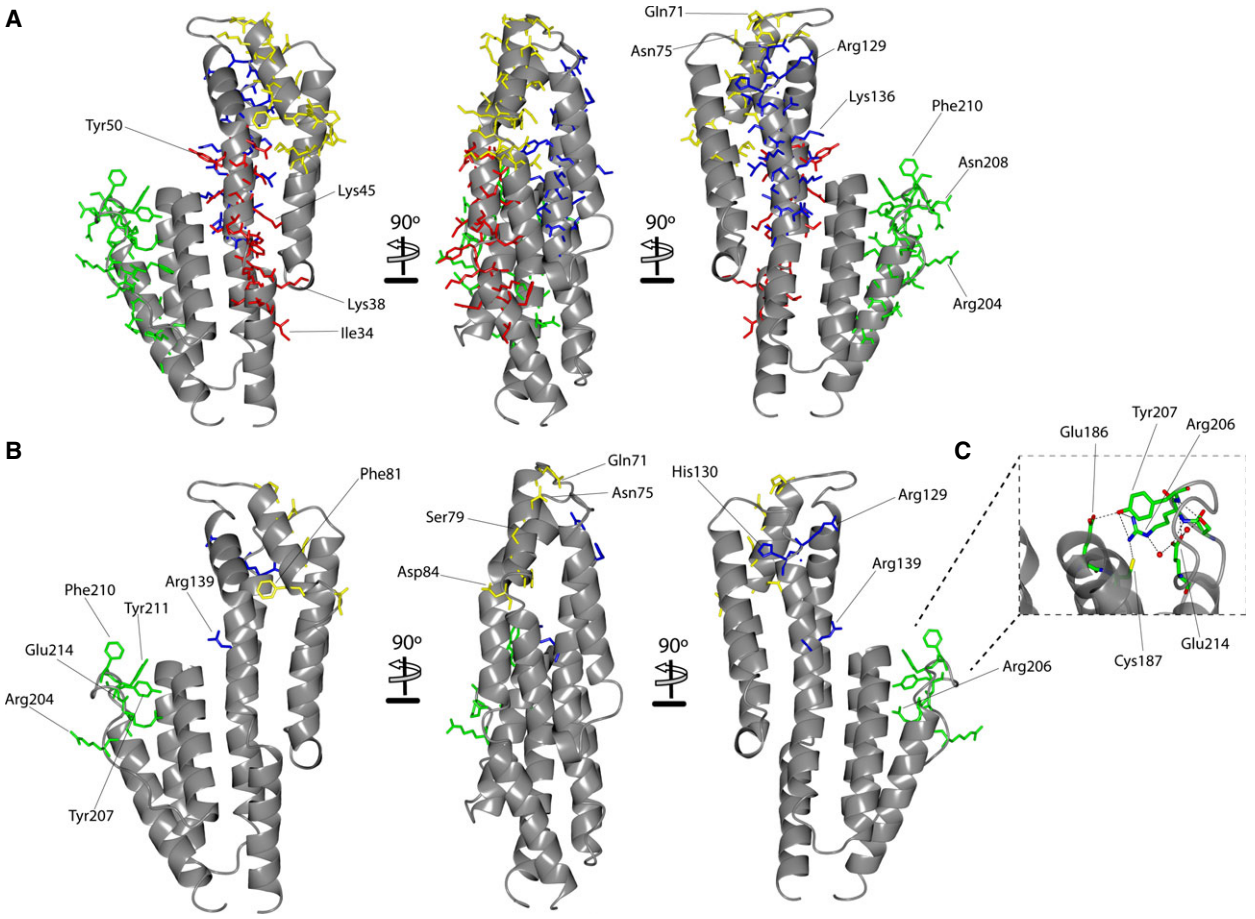

Fig. 2. Potentially relevant residues for CFH/CFHL-1 binding in CspZ. (A) Regions involved in CFH/CFHL-1 binding as found by Hartmann *et al.* using PEPSPOT analysis. The molecule is presented in three different orientations, rotated by 90° in the vertical plane. The residues belonging to the first region are colored red, the second region yellow, the third region blue and the fourth region green. (B) The residues involved in CFH/CFHL-1 binding in CspZ as found by Siegel *et al.* using the alanine-scanning approach. The coloring and orientation of the molecule are the same as in (A). (C) Location of residues Arg206, Tyr207 and Glu214 that showed significantly impaired binding to CFH/CFHL-1 after substitution to alanine reveals their importance in fold stabilization by hydrogen bonding with nearby residues.

Differences between homologous proteins

Some CspZ orthologs from *B. burgdorferi* and CspZ proteins from *B. garinii* and *B. afzelii* do not bind CFH/CFHL-1 [17,24]. One reason why CspZ proteins from *B. garinii* and *B. afzelii* do not bind CFH/CFHL-1 was thought to be associated with the presence of approximately 40 additional residues at the N-terminus. However, later investigations showed that the N-terminal extension does not influence ligand binding in CspZ [17]. The protein sequences from *B. burgdorferi* B31 and *B. burgdorferi* B408, which can bind CFH/CFHL-1, were compared with the sequences of *B. garinii* IP89, *B. burgdorferi* BL224 and *B. burgdorferi* B379, which cannot bind CFH/CFHL-1. The primary sequence diversity between these two protein groups may indicate that the loop region between α-helices αB and αC (residues 69–71 in *B. burgdorferi* B31) is involved in the binding of CspZ to CFH/CFHL-1, which is in agreement with previously reported results [24] (Fig. 3). However, in addition to the differences observed in the loop region, several differences appear in the sequence alignment between the homologous proteins. Nearly all of the divergent residues occur in only one or two of the three homologous proteins that do not bind CFH/CFHL-1 that were used in the sequence alignment, while the other homolog has the same residue as the proteins that are capable of CFH/CFHL-1 binding. Although this result indicates that those residues are not related to CFH/CFHL-1 binding, residues Ser72, Asp133 and Lys136, which were found to be different between the homologous proteins, were also checked by alanine mutagenesis by Siegel *et al.*, and the results of that study further validate that these residues are not involved in CFH/CFHL-1 binding [25].

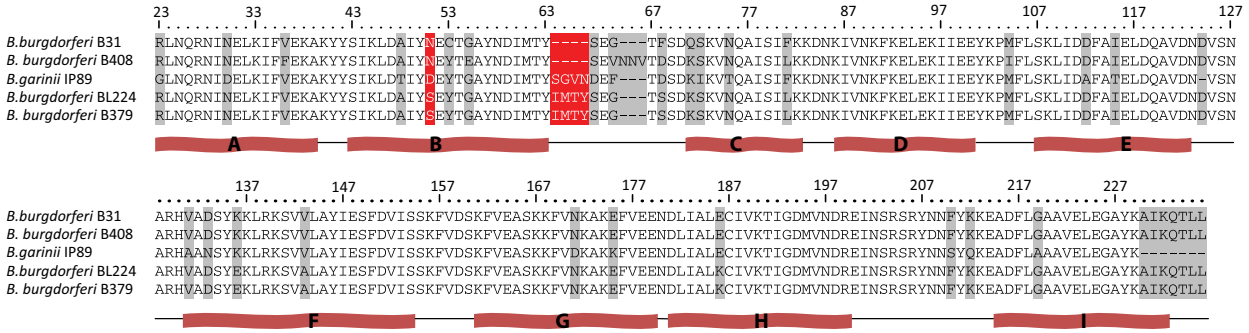

Fig. 3. Sequence alignments of the orthologous CspZ proteins that can bind to CFH/CFHL-1 (*B. burgdorferi* B31 and *B. burgdorferi* B408) and those that cannot bind to CFH/CFHL-1 (*B. garinii* IP89, *B. burgdorferi* BL224 and *B. burgdorferi* B379). The residues that are different for at least one orthologous protein but are not considered to be involved in CFH/CFHL-1 binding are colored grey. The residues proposed to be involved in CFH/CFHL-1 binding are colored red. The numbering above the alignment corresponds to the CspZ residue numbers from *B. burgdorferi* B31. Secondary structure elements for CspZ from *B. burgdorferi* B31 are shown below the alignment, whereas the α-helices are indicated as cylinders and loop regions as lines.

Putative location of CFH/CFHL-1 binding site in CspZ

Apart from the loop region, which differs between homologous proteins and most probably forms a binding site in CspZ, residue Asn51 has also been found to vary among the homologous proteins that can bind CFH/CFHL-1 and those that cannot. Together, the loop region between α-helices αB and αC (residues 69–71 in *B. burgdorferi* B31), Asn51 and Phe210, Tyr211 and Arg204 from the fourth region (amino acids 202–226) are all located on one side of the CspZ molecule; jointly, these residues may form the binding site for CFH/CFHL-1 (Fig. 4).

Binding sites within CFH and CFHL-1 required for binding of CspZ

It has been shown that CspZ binds to CFH and CFHL-1 through SCR domains 6 and 7 of CHF/CHFL-1 [10]. Two peptides corresponding to SCR6 (residues 333–351) and SCR7 (residues 393–414) have been identified in CFH and CFHL-1 using a peptide library representing SCR domains 5–7 [16]. These peptides displayed strong CspZ binding ability and were thought to be located in the binding site of CFH and CFHL-1 [10,16]. However, in our studies the two synthetic peptides did not show any binding to CspZ as probed by one-dimensional proton NMR spectroscopy. No changes, such as peptide signal broadening or chemical shift perturbations, were observed in the spectra after CspZ addition that would indicate binding of the peptides to the protein. Additionally, co-crystallization of CspZ with the respective peptides produced structures that were identical to CspZ alone, with no apparent electron density for the peptides (data not shown). Thus, our data suggest that the binding of the SCR6 and SCR7 peptides to CspZ is either very weak (high micromolar or weaker) or that environmental conditions can significantly influence the binding affinity, since the ligand affinity blot was

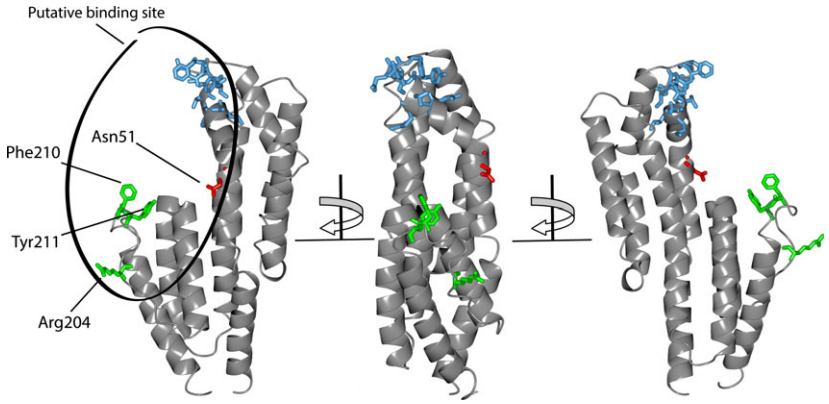

Fig. 4. Putative CspZ binding site. The potential binding site in CspZ includes residue Asn51 (red color), residues in the loop region between α-helices αB and αC (residues 59–68 colored in blue) and residues Arg204, Phe210 and Tyr211 (colored green) in the fourth region (residues 202–226).

carried out on a cellulose membrane while our studies were performed in solution.

Materials and methods

Cloning and expression of native and Se-Met-labeled CspZ

The CspZ gene from *B. burgdorferi* strain B31 was PCR-amplified from the genomic DNA. The predicted sequence coding for the signal peptide (residues 1–27) was excluded from the amplified product that was cloned into the pETm_11 expression vector (EMBL) encoding an N-terminal 6xHis tag followed by a TEV (tobacco etch virus) protease cleavage site. The plasmid was transformed into *Escherichia coli* RR1 cells and grown overnight at 37 °C on LB agar plates containing kanamycin. Colonies were inoculated into LB medium containing kanamycin and grown at 37 °C for another 24 h. Plasmid DNA was isolated from the resulting culture and verified by DNA sequencing. The plasmid was transformed in *E. coli* BL21 (DE3) cells and grown in modified 2xTYP medium [supplemented with kanamycin (10 $mg{\cdot}mL^{-1}$), 133 mM phosphate buffer (pH 7.4) and glucose (4 $g{\cdot}L^{-1}$)] with vigorous agitation at 25 °C until an A_{600} of 0.8–1.0. It was then induced with 0.2 mM isopropyl thio-β-D-galactoside and grown for an additional 16–20 h.

L-Selenomethionine (Se-Met) labeled protein was expressed in *E. coli* B834 (DE3). The cells were grown in modified 2xTYP medium [supplemented with 133 mM phosphate buffer (pH 7.4) and glucose (2 $g{\cdot}L^{-1}$)] until an A of 1.0 was reached. The cells were then centrifuged, and the resulting pellet was resuspended in 0501 medium (without methionine; Athena Enzyme Systems, Baltimore, USA) supplemented with 0502 medium (Athena Enzyme Systems) and glucose (5 $g{\cdot}L^{-1}$), and they were grown for an additional 2 h. Finally, 0.5 mM isopropyl thio-β-D-galactoside and a mixture of Se-Met:Met (5 : 1) was added, and the cultivation was continued for 20 h. The incorporation of Se-Met was verified by mass spectrometry as described previously [32].

Protein purification and 6xHis tag cleavage

The cells were lysed by sonication, and the cellular debris was removed by centrifugation. The supernatant was loaded onto a Ni-nitrilotriacetic acid agarose (Qiagen, Hilden, Germany) column and eluted with 300 mM imidazole, pH 7.5. The elution buffer was then exchanged for 20 mM Tris/HCl (pH 8.0) using an Amicon centrifugal filter unit (Millipore, Billerica, USA).

The 6xHis tag was removed by overnight incubation with TEV protease. The protease, the digested 6xHis tag and any remnants of uncleaved protein were subsequently removed using an additional round of Ni-nitrilotriacetic acid chromatography. The eluted protein fraction was concentrated to 9 $mg{\cdot}mL^{-1}$ with an Amicon centrifugal filter unit (Millipore). The purity of the CspZ was estimated by SDS/PAGE analysis to be > 95%.

Table 2. Data processing, refinement and validation statistics. Values in parentheses are for the highest resolution bin.

Data set	Se-Met	Native
Space group	P 2_12_12	P 2_12_12
Unit cell dimensions		
a (Å)	90.48	90.69
b (Å)	48.89	48.84
c (Å)	44.31	44.15
Wavelength (Å)	0.9805	0.9718
Resolution (Å)	33.23–2.10	45.34–1.83
Highest resolution bin (Å)	2.21–2.10	1.93–1.83
No. of reflections	166 726	64 751
No. of unique reflections	12 042	17 877
Completeness (%)	100.0 (100.0)	99.8 (100.0)
R_{merge}	0.10 (0.43)	0.04 (0.16)
I/σ (I)	15.4 (5.7)	19.1 (7.6)
Multiplicity	13.8 (13.9)	3.6 (3.7)
Refinement		
R_{work}	0.190 (0.222)	0.197 (0.220)
R_{free}	0.226 (0.323)	0.239 (0.249)
Average *B*-factor (Å^2)		
Overall	36.9	23.7
From Wilson plot	31.9	21.3
No. of atoms		
Protein	1664	1698
Water	77	107
RMSD from ideal		
Bond lengths (Å)	0.013	0.019
Bond angles (°)	1.435	1.840
Ramachandran outliers (%)		
Residues in most favored regions (%)	97.49	97.54
Residues in allowed regions (%)	2.51	2.46
Outliers (%)	0.00	0.00

NMR binding studies

NMR spectra were acquired at 298 K on a 600 MHz Varian Unity Inova spectrometer equipped with an HCN triple-resonance pulsed-field-gradient cold probe.

The proton 1D NMR spectra of each peptide in the absence and presence of the CspZ protein were compared to test the binding of the peptides (Metabion International AG, Munich, Germany). The protein and peptide stock solutions were prepared in 10 mM Tris/HCl (pH 8.0) buffer and then diluted in the same buffer containing 8% (v/v) D_2O for the binding studies. The NMR samples initially contained 90 μM peptide in a total volume of 250 μL, to which 37.5 μL of a 600 μM CspZ solution was added, yielding a peptide to protein concentration ratio of 1 : 1.

Crystallization of CspZ

The crystallization conditions for CspZ were determined by mixing 1 μL of protein with 1 μL of reservoir solution from commercially available screens (Molecular Dimensions Ltd, Newmarket, UK) using the sitting-drop vapor-diffusion method. The crystals were obtained in 30% PEG 2000 MME, 0.1 M Tris/HCl (pH 8.5) and 0.2 M sodium acetate. The Se-Met-labeled protein was crystallized in the same conditions. As a cryoprotectant, mother liquor with 25% glycerol was used in both cases, and the crystals were flash-frozen in liquid nitrogen.

Data collection and structure determination

The X-ray diffraction data for the native and Se-Met-labeled protein were collected at the MAX-lab beamlines I911-3 and I911-2 (Lund, Sweden). For the Se-Met-labeled CspZ protein, the structure was determined using the single-wavelength anomalous dispersion method. For the native CspZ protein, the structure was determined by molecular replacement using the structure of the Se-Met-labeled protein as the search model (PDB entry 4BG0). The reflections were indexed and scaled using MOSFLM [33] and SCALA [34] from the CCP4 suite [35]. The initial phases for Se-Met CspZ were obtained using SHELX C/D/E [36] and PHASER for the molecular replacement [37]. The initial protein model was built automatically in BUCCANEER [38], and minor rebuilding of the model was performed manually in COOT [39]. Water molecules were automatically selected in COOT and inspected manually. Crystallographic refinement was performed with REFMAC5 [40]. A summary of the data collection, refinement and validation statistics for CspZ is given in Table 2.

Accession numbers

The coordinates and structure factors for BBH06 (CspZ) have been deposited in the Protein Data Bank with the accession numbers 4BG0 and 4CBE.

Acknowledgements

This work was supported by ERAF grant 2010/0311/2DP/2.1.1.1.0/10/APIA/VIAA/022, ESF grant 1DP/1.1.1.2.0/09/APIA/VIAA/150 and the Latvian Council of Science No. 10.0029.3. We thank Dr Gunter Stier and Dr Huseyin Besir from the EMBL for providing the expression vector pETm-11. We acknowledge the staff at the MAX-lab synchrotron for their professional support during data collection.

Author contribution

Study conception and design, planning and performing of experiments, analyzing the data were done and the paper was also written by K.B. I.P. performed experiments. A.K. performed experiments. J.B. performed experiments. M.O. performed experiments. K.J. performed experiments. R.R. was the scientific advisor. K.T. performed experiments, analyzed data and revised the manuscript.

References

1 Barbour AG & Fish D (1993) The biological and social phenomenon of Lyme disease. *Science* **260**, 1610–1616.
2 Steere AC, Coburn J & Glickstein L (2004) The emergence of Lyme disease. *J Clin Invest* **113**, 1093–1101.
3 Stanek G, Wormser GP, Gray J & Strle F (2012) Lyme borreliosis. *Lancet* **379**, 461–473.
4 Zipfel PF, Skerka C, Hellwage J, Jokiranta ST, Meri S, Brade V, Kraiczy P, Noris M & Remuzzi G (2002) Factor H family proteins: on complement, microbes and human diseases. *Biochem Soc Trans* **30**, 971–978.
5 Zipfel PF & Skerka C (1999) FHL-1/reconectin: a human complement and immune regulator with cell-adhesive function. *Immunol Today* **20**, 135–140.
6 Zipfel PF, Jokiranta TS, Hellwage J, Koistinen V & Meri S (1999) The factor H protein family. *Immunopharmacology* **42**, 53–60.
7 Pangburn MK, Schreiber RD & Muller-Eberhard HJ (1977) Human complement C3b inactivator: isolation, characterization, and demonstration of an absolute requirement for the serum protein beta1H for cleavage of C3b and C4b in solution. *J Exp Med* **146**, 257–270.
8 Kuhn S, Skerka C & Zipfel PF (1995) Mapping of the complement regulatory domains in the human factor H-like protein 1 and in factor H1. *J Immunol* **155**, 5663–5670.
9 Schneider MC, Prosser BE, Caesar JJ, Kugelberg E, Li S, Zhang Q, Quoraishi S, Lovett JE, Deane JE, Sim RB *et al.* (2009) *Neisseria meningitidis* recruits factor H using protein mimicry of host carbohydrates. *Nature* **458**, 890–893.
10 Kraiczy P, Skerka C, Brade V & Zipfel PF (2001) Further characterization of complement regulator-acquiring surface proteins of *Borrelia burgdorferi*. *Infect Immun* **69**, 7800–7809.
11 Kraiczy P, Skerka C, Kirschfink M, Brade V & Zipfel PF (2001) Immune evasion of *Borrelia burgdorferi* by acquisition of human complement regulators FHL-1/reconectin and factor H. *Eur J Immunol* **31**, 1674–1684.
12 Kraiczy P & Stevenson B (2013) Complement regulator-acquiring surface proteins of *Borrelia burgdorferi*: structure, function and regulation of gene expression. *Ticks Tick Borne Dis* **4**, 26–34.
13 Kraiczy P, Hellwage J, Skerka C, Becker H, Kirschfink M, Simon MM, Brade V, Zipfel PF &

Wallich R (2004) Complement resistance of *Borrelia burgdorferi* correlates with the expression of BbCRASP-1, a novel linear plasmid-encoded surface protein that interacts with human factor H and FHL-1 and is unrelated to Erp proteins. *J Biol Chem* **279**, 2421–2429.

14 Cordes FS, Roversi P, Kraiczy P, Simon MM, Brade V, Jahraus O, Wallis R, Skerka C, Zipfel PF, Wallich R *et al*. (2005) A novel fold for the factor H-binding protein BbCRASP-1 of *Borrelia burgdorferi*. *Nat Struct Mol Biol* **12**, 276–277.

15 Caesar JJ, Wallich R, Kraiczy P, Zipfel PF & Lea SM (2013) Further structural insights into the binding of complement factor H by complement regulator-acquiring surface protein 1 (CspA) of *Borrelia burgdorferi*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **69**, 629–633.

16 Hartmann K, Corvey C, Skerka C, Kirschfink M, Karas M, Brade V, Miller JC, Stevenson B, Wallich R, Zipfel PF *et al*. (2006) Functional characterization of BbCRASP-2, a distinct outer membrane protein of *Borrelia burgdorferi* that binds host complement regulators factor H and FHL-1. *Mol Microbiol* **61**, 1220–1236.

17 Rogers EA & Marconi RT (2007) Delineation of species-specific binding properties of the CspZ protein (BBH06) of Lyme disease spirochetes: evidence for new contributions to the pathogenesis of *Borrelia* spp. *Infect Immun* **75**, 5272–5281.

18 Kraiczy P, Hartmann K, Hellwage J, Skerka C, Kirschfink M, Brade V, Zipfel PF, Wallich R & Stevenson B (2004) Immunological characterization of the complement regulator factor H-binding CRASP and Erp proteins of *Borrelia burgdorferi*. *Int J Med Microbiol* **293** (Suppl 37), 152–157.

19 Kraiczy P, Hellwage J, Skerka C, Kirschfink M, Brade V, Zipfel PF & Wallich R (2003) Immune evasion of *Borrelia burgdorferi*: mapping of a complement-inhibitor factor H-binding site of BbCRASP-3, a novel member of the Erp protein family. *Eur J Immunol* **33**, 697–707.

20 Hammerschmidt C, Hallstrom T, Skerka C, Wallich R, Stevenson B, Zipfel PF & Kraiczy P (2012) Contribution of the infection-associated complement regulator-acquiring surface protein 4 (ErpC) to complement resistance of *Borrelia burgdorferi*. *Clin Dev Immunol* **2012**, 349657.

21 Bhattacharjee A, Oeemig JS, Kolodziejczyk R, Meri T, Kajander T, Lehtinen MJ, Iwai H, Jokiranta TS & Goldman A (2013) Structural basis for complement evasion by Lyme disease pathogen *Borrelia burgdorferi*. *J Biol Chem*, **288**, 18685–18695.

22 Holm L, Kaariainen S, Rosenstrom P & Schenkel A (2008) Searching protein structure databases with DaliLite v. 3. *Bioinformatics* **24**, 2780–2781.

23 Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256–2268.

24 Rogers EA, Abdunnur SV, McDowell JV & Marconi RT (2009) Comparative analysis of the properties and ligand binding characteristics of CspZ, a factor H binding protein, derived from *Borrelia burgdorferi* isolates of human origin. *Infect Immun* **77**, 4396–4405.

25 Siegel C, Schreiber J, Haupt K, Skerka C, Brade V, Simon MM, Stevenson B, Wallich R, Zipfel PF & Kraiczy P (2008) Deciphering the ligand-binding sites in the *Borrelia burgdorferi* complement regulator-acquiring surface protein 2 required for interactions with the human immune regulators factor H and factor H-like protein 1. *J Biol Chem* **283**, 34855–34863.

26 Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S *et al*. (2012) InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res* **40**, D306–D312.

27 Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.

28 Cary LA & Guan JL (1999) Focal adhesion kinase in integrin-mediated signaling. *Front Biosci* **4**, D102–D113.

29 Parsons JT, Martin KH, Slack JK, Taylor JM & Weed SA (2000) Focal adhesion kinase: a regulator of focal adhesion dynamics and cell movement. *Oncogene* **19**, 5606–5613.

30 Hallstrom T, Haupt K, Kraiczy P, Hortschansky P, Wallich R, Skerka C & Zipfel PF (2010) Complement regulator-acquiring surface protein 1 of *Borrelia burgdorferi* binds to human bone morphogenic protein 2, several extracellular matrix proteins, and plasminogen. *J Infect Dis* **202**, 490–498.

31 Funahashi J, Takano K, Yamagata Y & Yutani K (2000) Role of surface hydrophobic residues in the conformational stability of human lysozyme at three different positions. *Biochemistry* **39**, 14448–14456.

32 Brangulis K, Tars K, Petrovskis I, Kazaks A, Ranka R & Baumanis V (2013) Structure of an outer surface lipoprotein BBA64 from the Lyme disease agent *Borrelia burgdorferi* which is critical to ensure infection after a tick bite. *Acta Crystallogr D Biol Crystallogr* **69**, 1099–1107.

33 Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**, 48–57.

34 Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72–82.

35 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie

AG, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242.
36 Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* **64**, 112–122.
37 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
38 Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* **62**, 1002–1011.
39 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
40 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.
41 Modzelewska K, Newman LP, Desai R & Keely PJ (2006) Ack1 mediates Cdc42-dependent cell migration and signaling to p130Cas. *J Biol Chem* **281**, 37527–37535.
42 Gates J & Peifer M (2005) Can 1000 reviews be wrong? Actin, alpha-catenin, and adherens junctions. *Cell* **123**, 769–772.
43 Benjamin JM, Kwiatkowski AV, Yang C, Korobova F, Pokutta S, Svitkina T, Weis WI & Nelson WJ (2010) AlphaE-catenin regulates actin dynamics independently of cadherin-mediated cell–cell adhesion. *J Cell Biol* **189**, 339–352.