Journal of Theoretical Biology 462 (2019) 97–108

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtb

Predicting the ligand-binding properties of *Borrelia burgdorferi* s.s. Bmp proteins in light of the conserved features of related *Borrelia* proteins

Mia Åstrand[a], Julia Cuellar[b,c], Jukka Hytönen[b], Tiina A. Salminen[a,*]

[a] Structural Bioinformatics Laboratory, Biochemistry, Faculty of Science and Engineering, Åbo Akademi University, Tykistökatu 6 A, Turku FI-20520, Finland
[b] Institute of Biomedicine, Faculty of Medicine, University of Turku, Turku, Finland
[c] Turku Doctoral Programme for Molecular Medicine, University of Turku, Turku, Finland

ARTICLE INFO

Article history:
Received 30 May 2018
Revised 29 October 2018
Accepted 5 November 2018
Available online 9 November 2018

Keywords:
Homology modeling
Phylogenetics
Lyme borreliosis
Relapsing fever
Substrate-binding proteins

ABSTRACT

Bacteria of the genus *Borrelia* cause vector-borne infections like the most important hard tick-borne disease in the northern hemisphere, Lyme borreliosis (LB), and soft tick or louse transmitted relapsing fevers (RF), prevalent in temperate and tropical areas. *Borrelia burgdorferi* sensu lato (s.l.) includes several genospecies and causes LB in humans. In infected patients, *Borrelia burgdorferi* sensu stricto (s.s.) expresses the BmpA, BmpB, BmpC and BmpD proteins. The role of these proteins in the pathogenesis of LB remains incompletely characterized, but they are, however, closely related to *Treponema pallidum* PnrA (Purine nucleoside receptor A), a substrate-binding lipoprotein of the ATP-binding cassette (ABC) transporter family preferentially binding purine nucleosides. Based on 3D homology modeling, the Bmp proteins share the typical fold of the substrate-binding protein family and the ligand-binding properties of BmpA, BmpB and BmpD are highly similar, whereas those of BmpC differ markedly. Nevertheless, these residues are highly conserved within the genus *Borrelia* and the inferred phylogenetic tree also reveals that the RF *Borrelia* lack BmpB proteins but has an additional Bmp protein (BmpA2) missing in LB-causing *Borrelia burgdorferi* s.l. Our results indicate that the Bmp proteins could bind nucleosides, although BmpC might have a different ligand-binding specificity and, therefore, a distinct function. Furthermore, the work provides a means for classifying the Bmp proteins and supports further elucidation of the roles of these proteins.

1. Introduction

Bacterial species of the phylum *Spirochaetes* are the causative agents of many globally prevalent illnesses, such as the venereal disease syphilis, and the tick- and/or louse-borne diseases Lyme borreliosis (LB) and relapsing fever (RF) (Gupta et al., 2013). *Treponema pallidum* subspecies *pallidum* (*T. pallidum*) infection leads to syphilis, which affects about 36.4 million patients worldwide (Rowley et al., 2012). *Borrelia burgdorferi* sensu lato (*B. burgdorferi* s.l.) and the relapsing fever associated *Borrelia* species (RF *Borrelia*) are important human pathogens that cause LB and RF, respectively. *B. burgdorferi* s.l. is divided into several so called genospecies such as *B. burgdorferi* sensu stricto (s.s.), *B. afzelii* and *B. garinii*. The RF *Borrelia* also includes numerous species. There are approximately 3,00,000 LB infections in the United States as estimated by the Centers for Disease Control and Prevention (Kuehn, 2013), whereas in Europe, there are 60,000–2,00,000 LB infections annually (Hubálek, 2009; Sykes and Makiello, 2014). RF is endemic especially in Africa where up to 9% of all fever patients are infected with RF *Borrelia* (Cutler, 2015; Nordstrand et al., 2007). Currently in Europe, there are emerging cases of louse-borne RF among East-African asylum seekers (Cutler, 2016; Darcis et al., 2016; Hytönen et al., 2017).

T. pallidum, B. burgdorferi s.l. and RF *Borrelia* are gram-negative bacteria with a helical structure and a didermic cell envelope (Cutler, 2015; Radolf et al., 2012, 2016). The spirochetes have a limited metabolic and biosynthetic capacity (Cutler, 2015; Fraser et al., 1997, 1998). For example, *T. pallidum* and *B. burgdorferi* s.l. lack the enzymes for *de novo* purine biosynthesis, unlike the RF *Borrelia* (Pettersson et al., 2007). Yet, the spirochetes are able to persist in the vector and in various human tissues. Hence, the spirochetes must rely on the uptake of host nutrients, such as nucleosides, nucleotides, or nitrogenous bases to maintain their replicative capacity. However, there are only a few spirochetal transporters described. Purine nucleoside receptor A (PnrA) of *T. pallidum* was the first described transporter of nucleosides in

* Corresponding author.
E-mail address: tiina.salminen@abo.fi (T.A. Salminen).

https://doi.org/10.1016/j.jtbi.2018.11.004

spirochetes (Deka et al., 2006). It is characterized as a membrane lipoprotein functioning as an ATP-binding cassette (ABC)-type substrate-binding protein, which is a class of proteins usually associated with ABC-transporters in bacteria (Maqbool et al., 2015). The substrate-binding protein recognizes its substrate and delivers it to the membrane-bound subunits of an ABC-transporter (Maqbool et al., 2015). All substrate-binding proteins share a conserved structure consisting of two alpha/beta domains connected by a hinge region and they are classified into seven clusters based on structural similarities. The ligand-binding site is located in the region between the two domains (Scheepers et al., 2016).

The basic membrane proteins (Bmp) of *B. burgdorferi* s.s. are closely related to the *T. pallidum* PnrA protein (Ramamoorthy et al., 1996). The *bmp* genes are located in tandem on the *B. burgdorferi* s.s. chromosome in the following order *bmpD, bmpC, bmpA, bmpB* (Gorbacheva et al., 2000; Ramamoorthy et al., 1996). The BmpB, BmpC and BmpD proteins are located on the inner membrane facing towards the periplasmic space (Dowdell et al., 2017) whereas BmpA, is reported to locate on the outer membrane of the bacteria (Bryksin et al., 2010). Antibodies against BmpA are detected in Lyme borreliosis patients (Simpson et al., 1990) and in mice infected with *B. burgdorferi* s.s. (Simpson et al., 1991) and, thus, the presence of antibodies against BmpA is considered as an indication of *B. burgdorferi* s.s. infection (Verma et al., 2009). Antibodies against BmpB and BmpD have also been found in Lyme borreliosis patients (Bryksin et al., 2005), and BmpC has been shown to be continuously expressed during infection in mice (Liang et al., 2002). All the *bmp* genes are transcribed *in vitro* but in varying degrees, *bmpA* being transcribed at the highest level, *bmpC* at the lowest and *bmpB* and *bmpD* at intermediate levels (Dobrikova et al., 2001). *B. burgdorferi* s.l. have to survive in both tick and vertebrate hosts, and therefore, its protein expression differs depending on the environment (Hovius et al., 2007). However, to the best of our knowledge it is not known if the Bmp proteins are expressed in ticks.

The exact function of the Bmp proteins is not fully characterized. Based on Verma et al. (2009), all of the *B. burgdorferi* s.s. Bmp proteins are adhesion proteins binding to laminin *in vitro*. The outer surface localization of BmpA (Bryksin et al., 2010) and the fact that *B. burgdorferi* s.s. upregulates *bmpA* expression specifically in the joint tissue, indeed, support the adhesive and possible proinflammatory function of BmpA (Pal et al., 2008; Yang et al., 2008; Zhao et al., 2017). However, the localization of BmpB, C and D in the inner membrane on the periplasmic side (Dowdell et al., 2017) suggests a non-adhesin function for these proteins. The Bmp proteins are homologous to the PnrA protein in *T. pallidum* (Simpson et al., 1994), which is located on the periplasmic side of the inner membrane and functions as a purine nucleoside transporter (Deka et al., 2006). Altogether, this indicates that the Bmp proteins might also be purine nucleoside transporters, meaning that they would, in fact, be multifunctional proteins, which has also been described before concerning some other *B. burgdorferi* s.l. proteins (reviewed in Caine and Coburn, 2016). The conflicting reports on the location and function of the four Bmp proteins raise many questions regarding these proteins and further studies are needed to conclusively determine their exact location and ascertain whether or not they are multifunctional proteins.

Although the Bmp proteins have been well characterized at the sequence level, their ligand-binding properties are still incompletely understood. To shed light on the 3D structures of the Bmp proteins, we have created 3D structural models and analyzed their ligand-binding properties. Furthermore, to study the molecular evolution and to identify differences and similarities between the Bmp proteins, a phylogenetic tree was constructed of related *Borrelia* proteins.

2. Methods

2.1. Sequence and structure analysis

The *B. burgdorferi* s.s. BmpA, BmpB, BmpC, BmpD and Exported protein (ExP) sequences were obtained from UniProt Knowledgebase (UniProtKB IDs: Q45010, Q45011, P0CL65, P0CL55, O51298 (Bateman et al., 2017). To search for similar crystal structures, which could be used as structural templates for the modeling, a BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1990) search against the Protein Data Bank (PDB) (Berman et al., 2000) was performed. The *B. burgdorferi* s.s. Bmp protein sequences were used as query. The BLAST searches found three similar protein structures, *T. pallidum* PnrA (PDB ID: 2fqw, (Deka et al., 2006), *T. pallidum* RfuA (PDB ID: 4iil, (Deka et al., 2013) and *Aeropyrum pernix* (*A. pernix*) K1 solute-binding protein (PDB ID: 4pev, to be published), with *E*-values below 9×10^{-10} (*E*-values < 0.001 are statistically significant (Pearson, 2013)). A structure-based multiple sequence alignment was created with the program VERTAA (Johnson and Lehtonen, 2000) by superimposing the abovementioned crystal structures.

The UniProtKB database (Bateman et al., 2017) was searched for Bmp and ExP sequences in *B. burgdorferi* s.l. (*B. burgdorferi* s.s. ATCC 35210 / B31 / CIP 102532 / DSM 4680, *B. burgdorferi* s.s. ZS7, *B. burgdorferi* s.s. 64b, *B. afzelii* PKo, *B. mayonii* MN14-1539, *B. bissettii* CO275, *B. bissettii* DN127, *B. garinii* PBr) and RF *Borrelia* (*B. turicatae* BTE5EL, *B. turicatae* 91E135, *B. hermsii* DAH-2E7, *B. hermsii* MTW, *B. duttonii* CR2A, *B. duttonii* Ly, *B. recurrentis* A1, *B. crocidurae* DOU, *B. crocidurae* Achema). Full length sequences from species with known strains were selected to create a multiple sequence alignment. To increase the reliability of the alignment, the sequences were aligned to the previously generated pre-aligned structure-based alignment using MALIGN in the Bodil modeling environment (Lehtonen et al., 2004). ESPript 3.0 (Robert and Gouet, 2014) was used to prepare pictures of the alignment.

Secondary structure predictions for the Bmp sequences were made with PSIPred (Jones, 1999), and LipoP (Juncker and Willenbrock, 2003) was used to predict lipoprotein signal peptides. The InterPro sequence analysis tool (Finn et al., 2017) was used to classify the proteins into protein families and to analyze the domain architecture of the proteins. The interactions between the ligand and the proteins were analyzed by visual inspection in PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC), by the Plip protein-ligand interaction server (Salentin et al., 2015) and the PDBePISA tool (Krissinel and Henrick, 2007).

2.2. 3D modeling

3D models were created using the generated multiple sequence alignment (Fig. 1) from which all sequences except the relevant *B. burgdorferi* s.s. Bmp protein and the template sequences had been removed. The crystal structure of *T. pallidum* PnrA (PDB ID: 2fqw, (Deka et al., 2006)), which has the highest sequence identity to the Bmp proteins, was chosen for the structural template. Ten models were created for each Bmp protein with the program MODELLER (Šali and Blundell, 1993) and the one with the lowest energy according to the MODELLER objective function was used for further analyses. The quality evaluation servers ProSA (Sippl, 1993; Wiederstein and Sippl, 2007), ProQ (Wallner and Elofsson, 2003) and MODFOLD (Maghrabi and Mcguffin, 2017) were used to assess the quality of the models. The models were also visually inspected and evaluated by superimposition with the structural template using the program VERTAA in the Bodil modeling environment (Johnson and Lehtonen, 2000). PyMOL was used for visualization and for preparing pictures (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC). The DSSP andStride

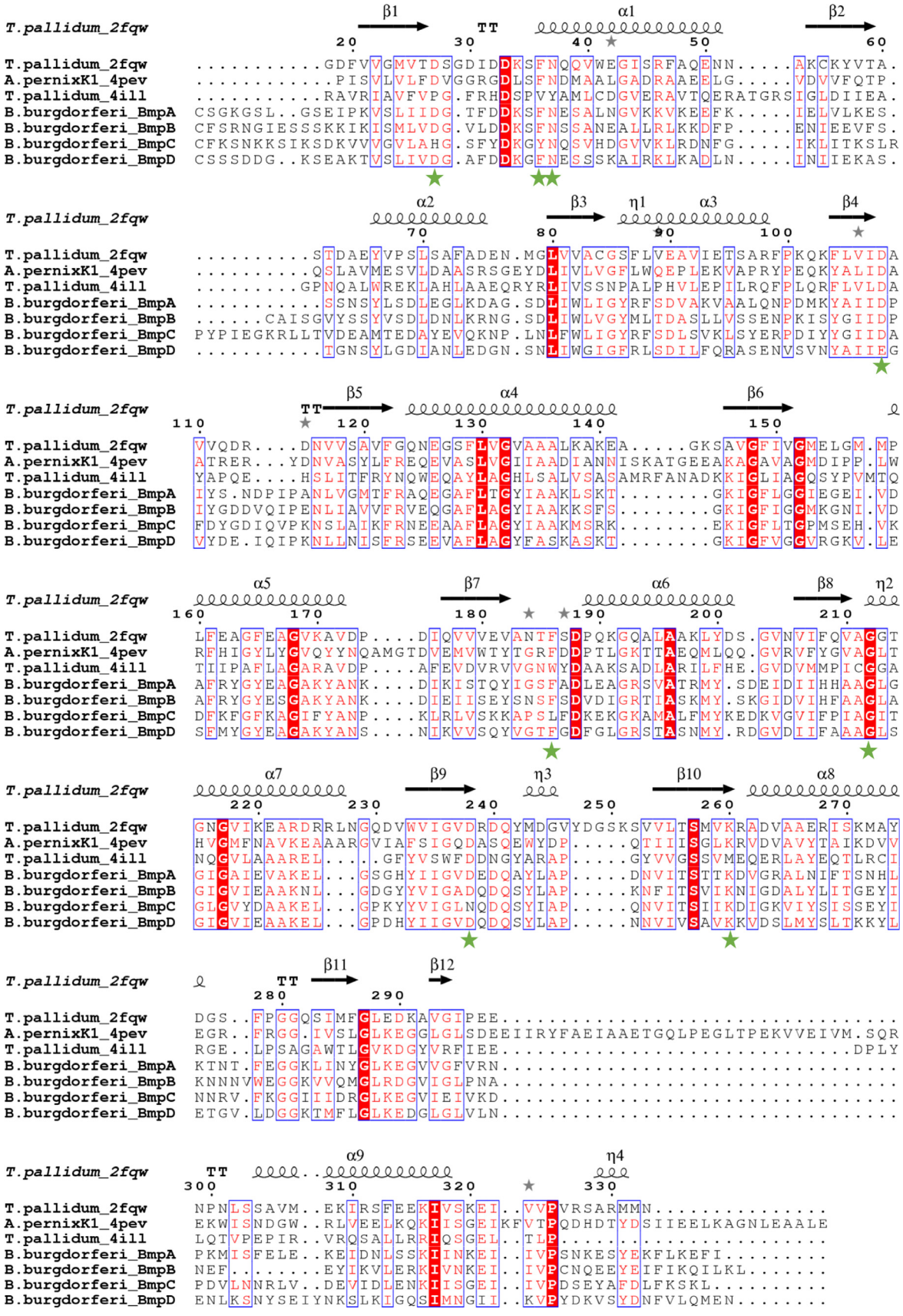


Fig. 1. Multiple sequence alignment used for modeling. The Bmp proteins aligned to the structure-based alignment of the templates *T. pallidum* PnrA (PDB ID 2fqw, (Deka et al., 2006)), *A. pernix* K1 solute-binding protein (PDB ID: 4pev, to be published) and *T. pallidum* RfuA (PDB ID: 4iil, (Deka et al., 2013)). The secondary structure of *T. pallidum* PnrA (PDB ID: 2fqw) is shown above the alignment. Conserved residues are shown with red background and similar residues are shown in red and boxed. Active site residues are marked with green stars.

Table 1
Sequence identities (%) between the *B. burgdorferi* s.s. Bmp proteins and the structural templates *T. pallidum* PnrA (PDB ID: 2fqw), *A. pernix* K1 solute-binding protein (PDB ID: 4pev) and *T. pallidum* RfuA (PDB ID: 4iil). The sequence identity between each Bmp protein and *T. pallidum* PnrA, which was used as the structural template for modeling, is shown in bold.

Protein	*A. pernix* K1 solute-binding protein	*T. pallidum* RufA	*T. pallidum* PnrA	BmpD	BmpC	BmpB
BmpA	27.1	16.9	**28.2**	47.2	37.5	50.7
BmpB	25.8	18.8	**31.3**	44.0	37.8	
BmpC	23.8	17.5	**27.5**	35.5		
BmpD	27.3	18.5	**27.8**			
***T. pallidum* PnrA**	27.5	23.9				
***T. pallidum* RfuA**	16.6					

plug-in in PyMOL was used to correctly assign the secondary structures in the models.

2.3. Phylogenetic analysis

The multiple sequence alignment was used to construct a Maximum Likelihood (ML) tree in MEGA7 (Kumar et al., 2016). The "Find Best DNA/Protein Models" option in MEGA7 was used to determine the appropriate evolutionary model for our data set. The Le and Gascuel (2008) substitution matrix (LG) with gamma distributions and invariant sites (G + I) was determined to be the best model, and thus used for the ML analysis. Furthermore, complete deletion of gaps and missing data was carried out to exclude highly variable regions from analysis and bootstrapping (500 replications) was used to evaluate branch support. Bootstrapping is a commonly used method for evaluating the confidence of a phylogenetic tree (Felsenstein, 1985; Hillis and Bull, 2010) and 500 replicates is the default setting in MEGA7 (Kumar et al., 2016).

3. Results and discussion

3.1. Sequence analysis

The four Bmp sequences in *B. burgdorferi* s.s. have 339 (BmpA), 341 (BmpB, BmpD), and 353 (BmpC) amino acids with the pairwise sequence identities in the range of 37.2–50.4% (Table 1). According to InterPro (Finn et al., 2017), all the proteins contain the ABC transporter substrate-binding protein PnrA-like (IPR003760) domain and the Periplasmic binding protein-like I (IPR028082) domain (Fig. 2A). They were also predicted to have a signal peptide by the LipoP scan (Juncker and Willenbrock, 2003). A region of approximately 20 amino acids typically constitutes the signal sequence, which usually spans from the first amino acid (Met) until the first cysteine residue (Zückert, 2014). The signal peptide is cleaved from the protein, and the bacterial lipoproteins are attached to the membrane by a diacylglyceryl group, bound to the sulfhydryl group of the cysteine, as well as through an acyl group attached to the amino group (Buddelmeijer, 2015).

3.2. 3D modeling of the B. burgdorferi s.s. Bmp proteins

Homology modeling, based on the assumption that similar proteins share a common structure, is a widely used protein structure prediction method (França, 2015). Since protein structures are more conserved than protein sequences (Illergård et al., 2009) it is possible to model a protein based on a related 3D structure even though the sequence identity is low (Chung and Subbiah, 1996).

In this case, a multiple sequence alignment was done due to the relatively low sequence identities (27.5–31.3%) between the *B. burgdorferi* s.s. Bmp proteins and the structural template (Table 1). First, a structure-based alignment was made with the known crystal structures (*T. pallidum* PnrA (PDB ID: 2fqw, (Deka et al., 2006), *T. pallidum* RfuA (PDB ID: 4iil, (Deka et al., 2013)) and *A. pernix* K1 solute-binding protein (PDB ID: 4pev, to be published) in order to identify structurally conserved residues. By combining a structure-based alignment with a multiple sequence alignment, conserved amino acids can be identified more confidently (Fiser, 2010). For this reason, the multiple sequence alignment, containing the sequences of the studied *B. burgdorferi* s.s. strain, was aligned to the pre-aligned structure-based alignment. The models were made using the 1.71 Å crystal structure of *T. pallidum* PnrA (PDB ID: 2fqw) as a structural template. The ligand (inosine) and the water molecules of the crystal structure of *T. pallidum* PnrA were included in the modeling as the water molecules in the active site of *T. pallidum* PnrA are involved in ligand-binding.

The models were evaluated by visual inspection and through quality evaluation servers (Table 2). The overall folds of the Bmp models show high similarity to PnrA (root mean square deviations were 0.42, 0.54, 0.45 and 0.50 Å respectively). The ModFOLD (Maghrabi and Mcguffin, 2017) global model quality score for all the models was above 0.6 (score $>$ 0.4 indicates complete and confident models) and the *p*-value for each model was in the range of 10^{-8}–10^{-9} (values $\langle$ 0.001 indicates that there is less than a 1/1000 chance that the models are incorrect). The ProSA web (Wiederstein and Sippl, 2007) Z-scores gave values well within the range typically found for proteins of similar size (*Z*-score $<$ −7.0 for all the models). The LG scores (values $>$ 4 indicates extremely good models) and the MaxSub scores (values $>$ 0.1 indicates correct models) of ProQ (Wallner and Elofsson, 2003) further support the quality of the models. Based on the results from the visual inspection and the quality evaluation servers, we concluded that our models were of high quality.

3.3. Structural similarity of the bmp models to T. pallidum PnrA and other substrate-binding proteins

The *B. burgdorferi* s.s. Bmp models were superimposed on the *T. pallidum* PnrA structure to compare the overall structure and the ligand-binding interactions. The Bmp models consist of two domains linked by three connecting loops (Fig. 2B). This agrees with the typical fold found in cluster B-I type substrate-binding proteins (Scheepers et al., 2016), to which the nucleoside-binding *T. pallidum* PnrA also belongs. The B-I cluster contains proteins that mainly bind to sugars, sugar alcohols and autoinducer 2. The substrate-binding proteins are divided into seven structural clusters, and the B cluster is further divided into subclusters based on substrate preference, and whether or not one or two alpha helices disrupt the second loop connecting the two domains. In proteins belonging to subcluster B-I, like the Bmp proteins, the additional alpha helix is missing. Like in the other substrate-binding proteins, the active site is formed in the cleft between the two domains and residues from both domains participate in ligand binding. Substrate-binding proteins are characterized by a conformational change upon ligand binding. Without a ligand, the protein is in a more flexible open conformation where the two domains are separated (Berntsson et al., 2010). When a ligand binds, the

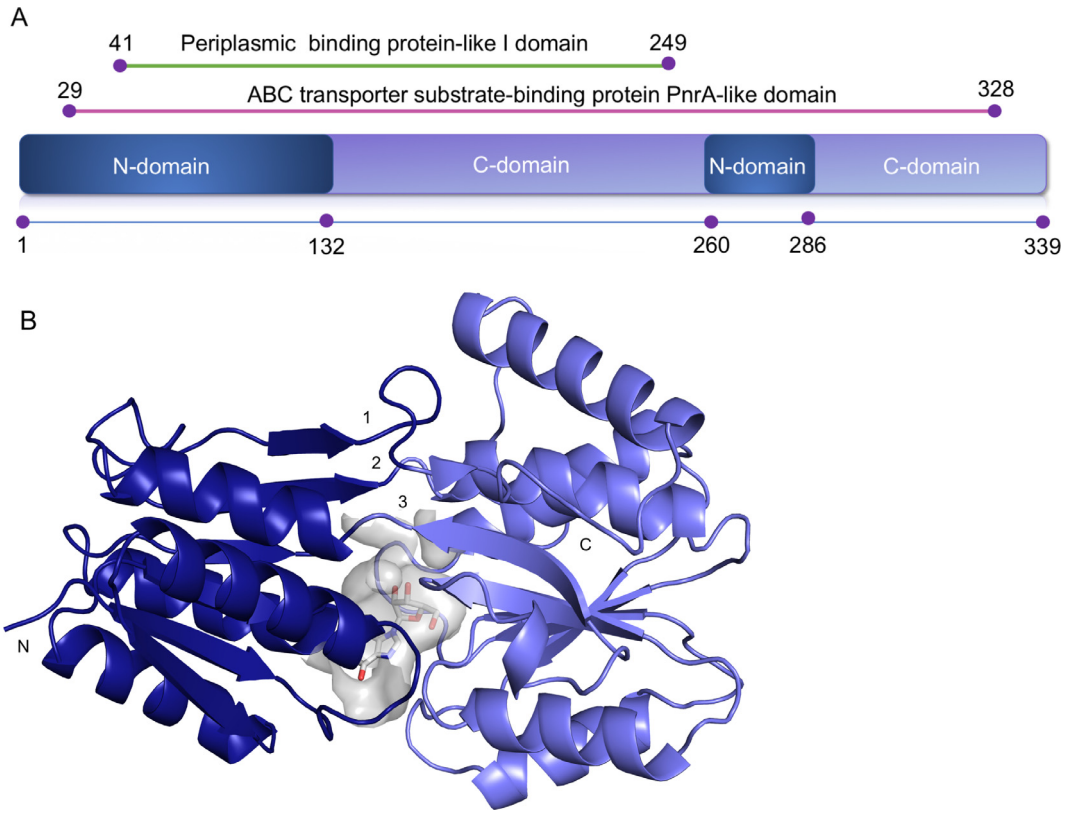


Fig. 2. A. Schematic picture of the BmpA protein with the location of the N-domain (residues 1–132, 260–286) and the C-domain (residues 133–259, 287–339) shown in the bar. The domains predicted by InterPro (Finn et al., 2017), IPR028082 Periplasmic binding protein-like I domain and IPR003760 ABC-transporter substrate-binding protein PnrA-like domain, are shown as lines above the bar. B. The BmpA model showing the two domains (N and C), the connecting loops (1–3) and the location of the active site pocket. The location of the active site pocket is marked in gray, and the ligand is shown as white sticks.

Table 2
Quality evaluation of the Bmp 3D models. Results from the quality evaluation servers (MODFOLD, ProQ and ProSA web) are shown.

Models	MODFOLD (*p*-value)	MODFOLD (global quality score)	ProQ (LGscore)	ProQ (MaxSub)	ProSA web (Z-score)
BmpA	1.68E−9	0.7509	4.150	0.485	−8.78
BmpB	7.263E−9	0.7171	4.097	0.488	−7.31
BmpC	2.735E−8	0.6864	4.148	0.510	−8.5
BmpD	9.788E−9	0.7102	4.140	0.454	−7.86

protein closes around the ligand in a process called the Venus fly trap mechanism (Mao et al., 1982). Thereafter, the substrate-binding protein binds to a membrane-bound transporter protein and delivers its ligand. Residues at both domains are important for binding to the transporter (Liu et al., 1999) and the substrate-binding proteins also play a role in stimulating the ATPase activity of the transporter (Davidson et al., 1992).

The two domains in the substrate-binding proteins are named after their location in the N-terminal (N domain) and the C-terminal (C domain) part. Both domains contain six parallel beta strands flanked by four alpha helices. The C domain also contains an additional alpha helix near the C-terminus. The last beta strand is very short and oriented antiparallel to the other beta strands. The structures of the Bmp models share the same fold as the *T. pallidum* PnrA, the differences are located mainly in variable surface loops. The *A. pernix* K1 solute-binding protein, on the other hand, contains four additional alpha helices in the C domain (not shown), whereas the N domain is more conserved and structurally similar to *T. pallidum* PnrA and the Bmp proteins. *T. pallidum* RfuA is another structurally similar substrate-binding protein, but in contrast to the Bmp proteins and the *T. pallidum* PnrA, it binds to riboflavin. Therefore, its binding site is quite different from those of the purine nucleoside-binding proteins.

3.4. Properties of the ligand-binding sites in the bmp models

The nucleoside inosine is bound to the crystal structure of *T. pallidum* PnrA (PDB ID: 2fqw, (Deka et al., 2006)) and was included in the modeling of the Bmp proteins to predict their nucleoside binding properties. In *T. pallidum* PnrA, the ribose part of the nucleoside forms hydrogen bonds with the side chains of Asp108, Asp238, Lys260 and with the main chain nitrogen of Gly212 (Fig. 3A). These interactions are conserved in the Bmp models with the following exceptions: Asp238 is replaced by an asparagine (Asn241) in BmpC and Asp108 is a glutamate (Glu98) in BmpD (Fig. 3E and F). In BmpC, Asn241 however forms the same hydrogen bonds to O3' and O5' as the aspartate in PnrA and the other models. Despite the longer side chain in glutamate, compared to aspartate, Glu98 in BmpD can still form the same interactions as the aspartate.

The residues binding to the purine base part of the ligand are conserved in BmpA, BmpB and BmpD. The aromatic side chains of Phe36 and Phe186 (PnrA numbering) form aromatic sandwich stacking with the imidazole and pyrimidine rings on the ligand. Asp27 is considered a crucial residue for purine binding in *T. pallidum* PnrA (Deka et al., 2006) and it forms hydrogen bonds with both N1 and O6 on the ligand (Fig. 3A). BmpC differs notably from

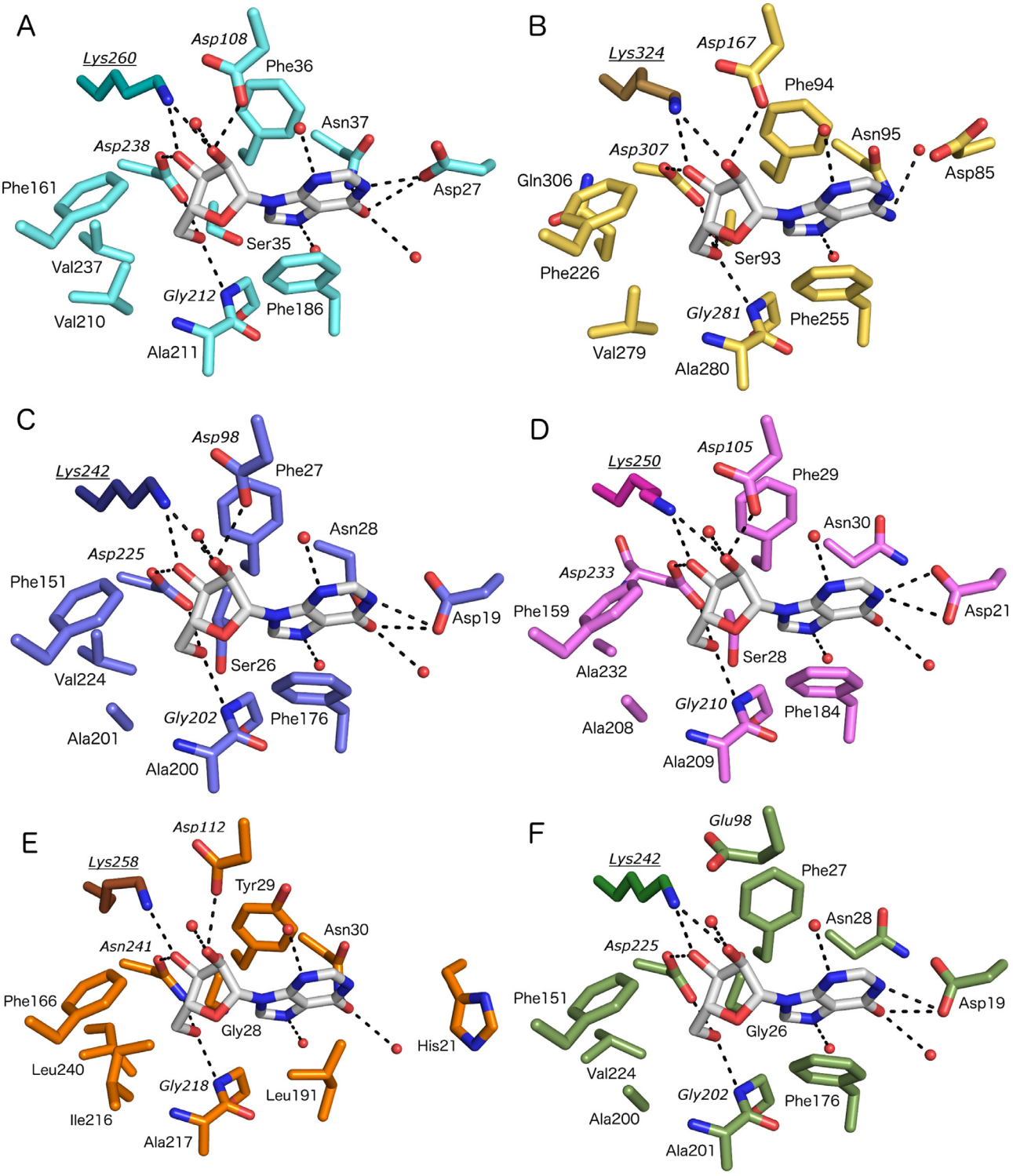


Fig. 3. Active site residues in the *T. pallidum* PnrA-inosine complex (A), the *A. pernix* K1 substrate-binding protein in complex with adenosine (B), and the Bmp models in complex with the potential ligand inosine, BmpA (C), BmpB (D), BmpC (E), BmpD (F). The completely conserved lysine is shown in a darker color with the residue name underlined, and residues binding to the ribose part are in italics.

the other Bmps in the residues binding to the purine base part. Tyr 29 and Leu191 in BmpC correspond to Phe36 and Phe186 in PnrA (Fig. 3E). Leu191 cannot form the aromatic stacking interaction with the purine base but the aromatic stacking interaction with Tyr29 is maintained. Furthermore, BmpC has a positively charged histidine (His21) instead of a negatively charged aspartate at Asp27 (PnrA numbering).

The ligand-binding residues in the *A. pernix* K1 solute-binding protein are also conserved forming the same interactions with the ligand as in *T. pallidum* PnrA (Table 3). Among the Bmp proteins, BmpA and BmpB share 50.7% sequence identity (Table 1) and their ligand-binding residues are totally conserved also with those of PnrA and the *A. pernix* K1 solute-binding protein (Fig. 3A–D). BmpD, with the Asp/Glu replacement in the ligand binding site

Table 3
A. Active site interactions in *T. pallidum* PnrA (PDB ID: 2fqw), *A. pernix* K1 solute-binding protein (PDB ID: 4pev) and in the Bmp models. The interactions were analyzed by visual inspection, by the Plip server (Salentin et al., 2015) and by PDBePISA (Krissinel and Henrick, 2007). Interactions noted only by visual inspection are marked with +. Interactions found by Plip, but not by PSDePISA, are marked with a * and interactions found by PDBePISA, but not by Plip, are marked with #. B. The naming convention of the ligands, inosine (left), adenosine (right), are shown.

A

***T. pallidum* PnrA**

Residue	Atom name	Inosine atom name	Distance (Å)	Interaction type
Gly212	N	O5'	2.85	Hydrogen bond
Lys260	NZ	O2'	2.92	Hydrogen bond #
Lys260	NZ	O3'	2.88	Hydrogen bond
Asp108	OD2	O2'	2.54	Hydrogen bond *
Asp238	OD2	O5'	2.73	Hydrogen bond
Asp238	OD2	O3'	2.63	Hydrogen bond *
Asp27	OD1	N1	2.6	Hydrogen bond +
Asp27	OD1	O6	3.4	Hydrogen bond +
Asn37	ND2	O6	3.05	Hydrogen bond
Phe186	-	Imidazole ring	3.60	Pi stacking *
Phe186	-	Pyrimidine ring	3.97	Pi stacking *
Asp27	OD1	O6	3.7/3.0	Water bridge *
Asp27	OD1	N1	3.7/4.0	Water bridge *
Ser28	OG	O6	2.7/3.7	Water bridge *

***A. pernix* K1 solute-binding protein**

Residue	Atom name	Adenosine atom name	Distance (Å)	Interaction type
Gly281	N	O5'	2.83	Hydrogen bond
Lys324	NZ	O2'	3.13	Hydrogen bond
Lys324	NZ	O3'	3.1	Hydrogen bond +
Asp167	OD2	O2'	2.9	Hydrogen bond +
Asp307	OD2	O5'	2.75	Hydrogen bond
Asp307	OD1	O3'	2.45	Hydrogen bond *
Asn95	OD1	N6	3.9	Hydrogen bond
Asn95	OD1	N6	3.49	Hydrogen bond *
Gly87	N	N6	3.77	Hydrogen bond *
Ser93	OG	O5'	3.44	Hydrogen bond *
Phe255	-	Imidazole ring	3.55	Pi stacking *
Phe255	-	Pyrimidine ring	3.74	Pi stacking *
Asn95	OD1	N6	2.6/3.2	Water bridge *
Val86	N	N1	3.8/3.8	Water bridge *
Asp91	OD2	N6	2.9/3.7	Water bridge *
Phe255	O	N6	3.5/3.7	Water bridge *

BmpA

Residue	Atom name	Inosine atom name	Distance (Å)	Interaction type
Gly202	N	O5'	2.84	Hydrogen bond
Lys242	NZ	O2'	3.20	Hydrogen bond
Lys242	NZ	O3'	3.25	Hydrogen bond
Asp98	OD2	O2'	2.81	Hydrogen bond *
Asp225	OD2	O5'	2.93	Hydrogen bond #
Asp225	OD2	O3'	2.8	Hydrogen bond +
Asp19	OD2	N1	3.1	Hydrogen bond +
Asp19	OD2	O6	3.2	Hydrogen bond +
Phe176	-	Imidazole ring	3.53	Pi stacking *
Phe176	-	Pyrimidine ring	3.86	Pi stacking *

BmpB

Residue	Atom name	Inosine atom name	Distance (Å)	Interaction type
Gly210	N	O5'	2.86	Hydrogen bond
Lys250	NZ	O2'	3.15	Hydrogen bond #
Lys250	NZ	O3'	3.00	Hydrogen bond #
Asp105	OD2	O2'	2.64	Hydrogen bond *
Asp233	OD2	O5'	2.87	Hydrogen bond #
Asp233	OD1	O3'	2.8	Hydrogen bond +
Asp21	OD1	O6	3.08	Hydrogen bond #
Asp21	OD1	N1	2.8	Hydrogen bond +
Asp21	OD2	N1	3.1	Hydrogen bond +
Phe184	-	Imidazole ring	3.65	Pi stacking *
Phe184	-	Pyrimidine ring	3.89	Pi stacking *
Lys250	NZ	O2'	3.18/3.89	Water bridge *

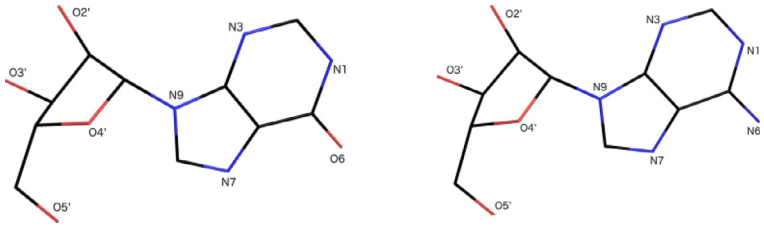
BmpC

Residue	Atom name	Inosine atom name	Distance (Å)	Interaction type
Gly218	N	O5'	2.84	Hydrogen bond
Lys258	NZ	O2'	3.57	Hydrogen bond #
Lys258	NZ	O3'	2.94	Hydrogen bond
Asn241	ND2	O5'	2.87	Hydrogen bond
Asn241	OD1	O3'	2.8	Hydrogen bond
Asp112	OD2	O2'	4.02	Hydrogen bond *
Asn30	ND2	O6	3.31	Hydrogen bond
Tyr29	-	Imidazole ring	3.84	Pi stacking *
Asp112	OD2	O2'	3.8/3.2	Water bridge *

BmpD

Residue	Atom name	Inosine atom name	Distance (Å)	Interaction type
Gly202	N	O5'	2.85	Hydrogen bond
Lys242	NZ	O2'	3.02	Hydrogen bond #
Lys242	NZ	O3'	2.93	Hydrogen bond
Lys146	NZ	O2'	3.63	Hydrogen bond
Asp225	OD2	O5'	2.7	Hydrogen bond +
Asp225	OD1	O3'	2.8	Hydrogen bond +
Asp19	OD2	N1	2.9	Hydrogen bond +
Asp19	OD2	O6	3.2	Hydrogen bond +
Phe27	-	Imidazole ring	5.19	Pi stacking *
Phe176	-	Imidazole ring	3.79	Pi stacking *
Phe176	-	Pyrimidine ring	3.61	Pi stacking *

B



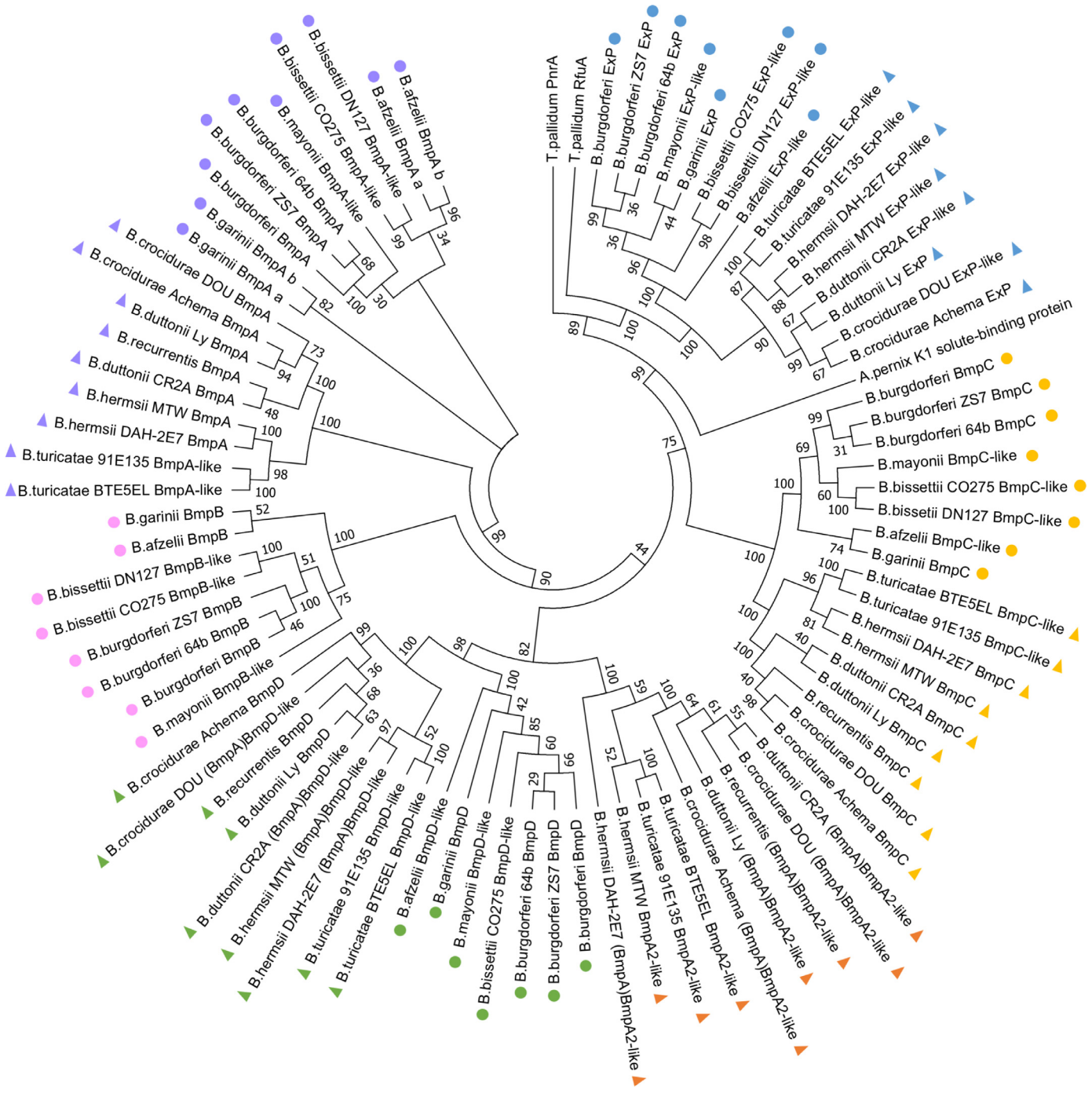


Fig. 4. Maximum likelihood tree of *Borrelia* Bmp and ExP proteins. The proteins form separate branches and are colored accordingly. *B. burgdorferi* s.l. are marked by circles and RF *Borrelia* by triangles. Branch support is shown as bootstrap values. The name of the strains is included in the name if there is more than one strain per species. See table S1 for complete information on included sequences. Proteins are named BmpX-like if the name of the protein is unspecified in the UniProtKB database (Bateman et al., 2017). If the UniProtKB protein name differs from how the protein is classified in the tree, the UniProt name is found in parenthesis.

(Asp108 in PnrA), is highly similar to BmpA and BmpB (47.2 and 44.0% sequence identities, respectively). BmpC, however, shares only 35.5–37.8% sequence identity with the other Bmps, and the residues binding to the base part of the nucleoside differs significantly from the other Bmps and the related proteins PnrA and *A. pernix* K1 solute-binding protein.

In summary, the residues involved in binding to the ribose moiety of the nucleoside in the studied proteins are highly conserved, or conservatively substituted (BmpC, BmpD) to maintain the interaction despite the residue changes. The residues binding to the purine part are notably less conserved in BmpC, which indicate that BmpC could prefer different nucleosides than the other Bmps, and thus, might have a distinctive function.

3.5. Molecular evolution of the Borrelia bmp proteins and organization of the corresponding genes

To clarify the evolutionary relationships between the *Borrelia* Bmp proteins and the related Exported proteins (ExP), a phylogenetic tree was inferred. Since many of the Bmp sequences in the UniProtKB database (Bateman et al., 2017) were not annotated as a specific Bmp protein we wanted to classify the proteins in

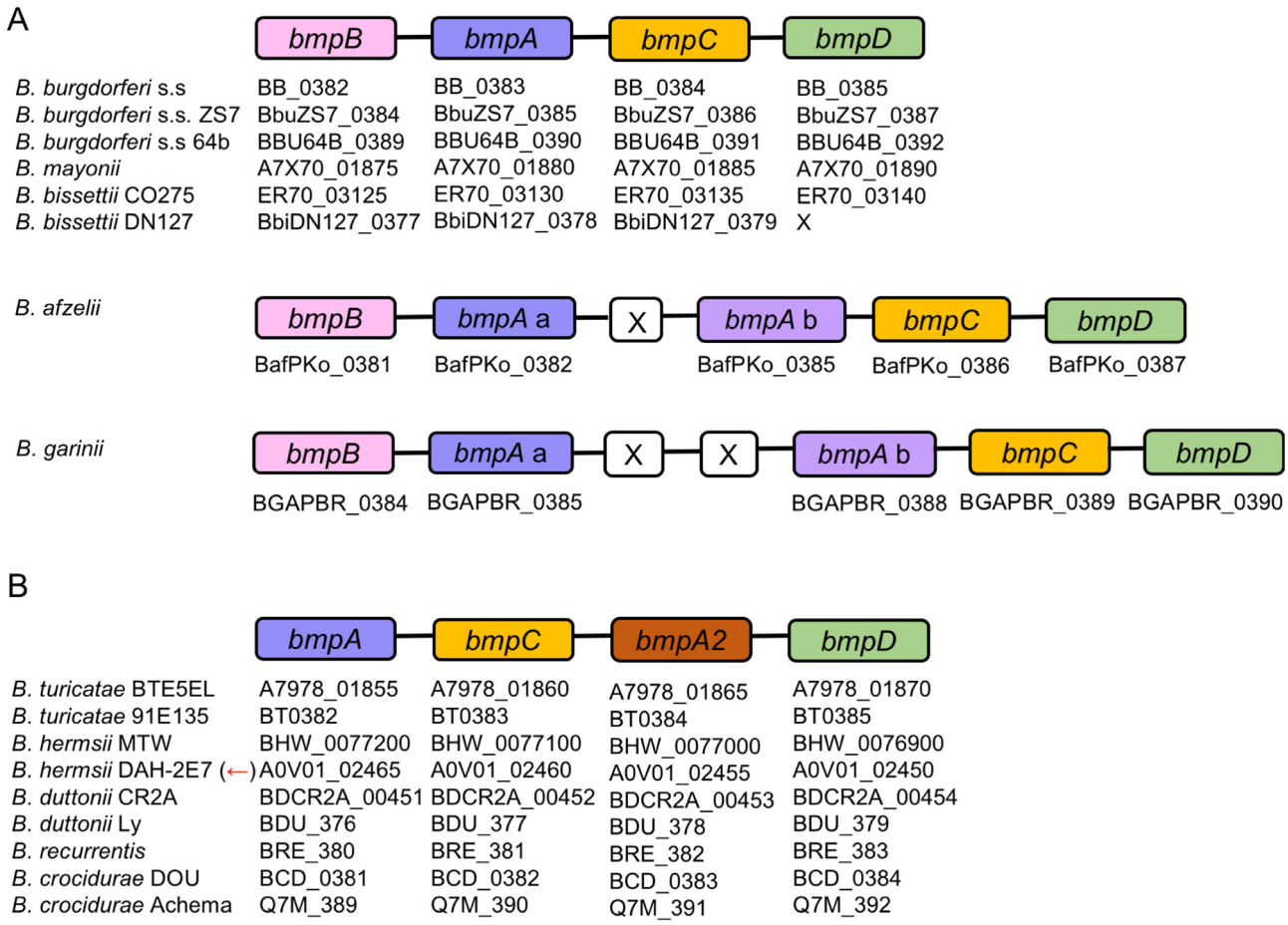


Fig. 5. Gene order of the *bmp* genes in *Borrelia* species (A = *B. burgdorferi* s.l., B = RF *Borrelia*), based on the phylogenetic analysis. X = fragment or missing sequence in the NCBI database (Agarwala et al., 2017). ← = the order of the genes is reversed compared to the others.

order to identify functional differences. As some of the ExP proteins were annotated as basic membrane proteins or nucleoside-binding proteins, they were included in the phylogenetic analysis. Furthermore, we were interested in ascertaining if the Bmp proteins would form distinct clades for the *B. burgdorferi* s.l. and RF *Borrelia*.

The ML method, which gives a tree with the highest likelihood of producing the original multiple sequence alignment, was used for the phylogenetic analysis. The ML-tree of the *Borrelia* Bmp and ExP proteins diverges into four main branches (Fig. 4). The first branch segregates into distinct branches for the ExP and the BmpC proteins. The second branch contains the BmpD and BmpA2 proteins. The BmpB proteins form the third main branch, and the last branch is made up of the BmpA proteins. Except for BmpB, the branches for each protein further diverges into separate sub branches for the *B. burgdorferi* s.l. and RF *Borrelia* proteins. For the BmpA proteins, however, the *B. burgdorferi* s.l. and RF *Borrelia* branches are completely separate and do not form a common branch before diverging into separate *B. burgdorferi* s.l. and RF *Borrelia* branches like the other proteins. All main branches are supported by high bootstrap values, except for the branch that splits into the ExP/BmpC and the BmpD/BmpA2 branches (bootstrap value 44). The low bootstrap value suggests that the sequences diverged during a relatively short evolutionary time frame (Soltis and Soltis, 2003, 2014). In agreement with the analyses done by Lescot et al. (2008), where no BmpB-like proteins were identified in the RF *Borrelia* (Lescot et al., 2008), the BmpB branch in the ML-tree contains only *B. burgdorferi* s.l. proteins.

Interestingly, the RF *Borrelia* contain an additional Bmp protein, which is not found in the *B. burgdorferi* s.l.. This additional protein has been labeled BmpA2 in Lescot et al. (2008), however, based on our phylogenetic tree, these proteins are found within the same branch as the BmpD proteins. In contrast to the BmpD proteins, however, the BmpA2 proteins have a Gly instead of Asn37 (PnrA numbering) in the ligand-binding site (Table S2). In fact, the Asn is conserved within all other studied Bmp proteins (except for the RF *Borrelia* BmpC proteins, which have Phe or Leu instead). As the glycine does not have any side chain, it cannot form the same interaction with O6 on the ligand as the Asn (Table S2).

The ExP proteins are, like the Bmp proteins, classified as ABC transporter substrate-binding proteins. In fact, the ExP have a higher sequence identity to the *T. pallidum* RfuA protein, which is an ABC-type riboflavin transporter in spirochetes, than to the *T. pallidum* PnrA protein (Deka et al., 2013). Consistently, the ExP proteins in the ML-tree are found within the same branch as the *T. pallidum* RfuA protein, indicating that the ExP proteins could similarly function as riboflavin-binding proteins.

The BmpC proteins share a more recent common ancestor with the ExP proteins than with the rest of the Bmp proteins. Despite this, the BmpC protein is most likely not a riboflavin-binding protein since its ligand-binding site is highly different from that of RfuA. The 3D structural model of BmpC (Fig. 3E) showed that the ligand-binding site in BmpC clearly differs from the rest of the Bmp proteins and, moreover, experiments have shown that the expression of BmpC is four or five times higher in infectious *B. burgdorferi* s.s. strain than in a non-infectious strain

whereas no such variations were found for the other Bmp proteins (Dobrikova et al., 2001). Furthermore, even within all the *Borrelia* BmpC proteins the residues binding to the ribose part of the nucleoside are conserved while the residues binding to the base moiety of the nucleoside vary considerably (Table S2). The position corresponding to His21 (*B. burgdorferi* s.s.; Fig. 3E) is either His or Asn in the *B. burgdorferi* s.l. BmpC proteins, but Pro or Ser in RF Borrelia. Similarly, Leu191 is Leu/Arg in the *B. burgdorferi* s.l. genospecies and Glu/Asp in the RF Borrelia. None of these residues could form stacking interactions with the base part of the nucleoside. Asn28 is conserved in all the *B. burgdorferi* s.l., while in the RF *Borrelia* the corresponding residue is Phe or Leu. Taken together, these differences strongly suggest that the function of the BmpC proteins might differ significantly from the other Bmp proteins. Furthermore, the residue variation between species suggest a possible functional difference also between the RF Borrelia and the *B. burgdorferi* s.l..

Studies have shown that all four Bmp proteins in *B. burgdorferi* s.s. are simultaneously expressed *in vitro*, but to varying degrees (Dobrikova et al., 2001). BmpA is expressed in significantly higher amounts than the other Bmp proteins, and BmpC in particular is expressed at very low levels. BmpA, BmpC and BmpD have their own transcriptional start sites, whereas BmpB is transcribed together with BmpA (Ramamoorthy et al., 2005). At the genome level, the order of the genes differs between the *B. burgdorferi* s.l. and the RF *Borrelia*. In both groups there are four different Bmp proteins in the same gene cluster (Fig. 5). The genomes of *B. afzelii* and *B. garinii* are exceptions to the rule, however, containing at least five different Bmp proteins. Possibly, duplications have occurred at some point during the evolution of these proteins. Both *B. afzelii* and *B. garinii* contain an additional BmpA protein. However, in contrast to the BmpA2 proteins found in the RF *Borrelia*, the extra BmpA proteins in the *B. burgdorferi* s.l. are very similar to (>94% sequence identity) and cluster together with the other BmpA protein in these genospecies. Therefore, we decided to label these proteins as BmpA a and BmpA b in order to distinguish them from the BmpA2 proteins found in the RF *Borrelia* (Fig. 5). Interestingly, the *B. afzelii* BmpA b contains a serine instead of Gly202 in BmpA and the *B. garinii* BmpA b protein contains an asparagine instead of Asp19 (Fig. 3C). Most likely, neither change has any impact on the ligand-binding properties. The side-chain nitrogen and oxygen of the asparagine can still form the same hydrogen bonds as the aspartate. Gly202 binds to the ribose part through the backbone nitrogen, which would not be affected by a change to serine. The larger side-chain of serine might, however, affect the binding pocket, depending on the position of the side-chain.

4. Conclusions

In this study, we have shown that the Bmp proteins likely consist of two alpha/beta domains connected by a hinge region with three connecting loops. This structure is typical of bacterial ABC transporter substrate-binding proteins (Berntsson et al., 2010). The structural similarity to *T. pallidum* PnrA indicates that the Bmp proteins belong to the B cluster in the classification of substrate-binding proteins (Scheepers et al., 2016). The similarities of the active sites in the Bmp proteins and the *T. pallidum* PnrA suggest that the Bmp proteins could also play a role in nucleoside transport. Access to nucleosides is vital for *B. burgdorferi* s.l. survival in the host, and since *B. burgdorferi* s.l. lacks the enzymes needed to synthesize purines *de novo* (Pettersson et al., 2007), the Bmp proteins could play a vital part in obtaining the purines needed to establish an infection in the host.

We also characterized the *Borrelia* Bmp proteins using a phylogenetic analysis, and this provides a means for classifying the Bmp proteins. The additional Bmp protein (BmpA2), found only in RF *Borrelia*, was observed to be more closely related to the BmpD proteins. The BmpB proteins were found only in the *B. burgdorferi* s.l., in agreement with Lescot et al. (2008). Moreover, the ExP proteins were shown to form a separate branch with the RfuA protein of *T. pallidum*, suggesting that they could function as riboflavin-binding proteins. Although the BmpC proteins are found within the same main branch as the ExP proteins, their ligand-binding site is clearly different from that of the riboflavin-binding protein RfuA.

Many questions regarding the function of the Bmp proteins still remain. For example, can the Bmp proteins bind to different purine nucleosides, or are they specific for a single one? Is it possible that the BmpC protein can bind to a non-purine, considering the differences in the ligand-binding site? Are all the Bmp proteins active at the same time in the host? Or are they activated at different times or in different tissues? Further experiments are needed to ascertain the function of the Bmp proteins and to answer these questions.

Acknowledgments

We thank Gabriela Guédez for helpful comments on the manuscript, the bioinformatics (J.V. Lehtonen), translational activities and structural biology infrastructure support from Biocenter Finland, and CSC IT Center for Science for laboratory and computational infrastructure support. Professor Mark Johnson is acknowledged for the excellent computing facilities at the Structural Bioinformatics Laboratory, Åbo Akademi University.

This work was supported by the National Doctoral Programme in Informational and Structural Biology(MÅ), Svenska Kulturfonden (MÅ), Orion Research Foundation (MÅ), Academy of Finland (JH), Jane and Aatos Erkko Foundation (JH), Sigrid Juselius Foundation (TAS) and the Tor, Joe, and Pentti Borg's Foundation (TAS).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

MÅ performed the research, analyzed the data and wrote the manuscript; JC participated in the data analysis and helped to draft the manuscript; JH participated in the design of the work and helped to draft the manuscript; TAS conceived and designed the work and wrote the manuscript.

All authors have approved the final manuscript.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jtbi.2018.11.004.

References

Agarwala, R., Barrett, T., Beck, J., Benson, D.A., Bollin, C., Bolton, E., Bourexis, D., Brister, J.R., Bryant, S.H., Canese, K., Charowhas, C., Clark, K., DiCuccio, M., Dondoshansky, I., Feolo, M., Funk, K., Geer, L.Y., Gorelenkov, V., Hlavina, W., Hoeppner, M., Holmes, B., Johnson, M., Khotomlianski, V., Kimchi, A., Kimelman, M., Kitts, P., Klimke, W., Krasnov, S., Kuznetsov, A., Landrum, M.J., Landsman, D., Lee, J.M., Lipman, D.J., Lu, Z., Madden, T.L., Madej, T., Marchler-Bauer, A., Karsch-Mizrachi, I., Murphy, T., Orris, R., Ostell, J., O'Sullivan, C., Palanigobu, V., Panchenko, A.R., Phan, L., Pruitt, K.D., Rodarmer, K., Rubinstein, W., Sayers, E.W., Schneider, V., Schoch, C.L., Schuler, G.D., Sherry, S.T., Sirotkin, K., Siyan, K., Slotta, D., Soboleva, A., Soussov, V., Starchenko, G., Tatusova, T.A., Todorov, K., Trawick, B.W., Vakatov, D., Wang, Y., Ward, M., Wilbur, W.J., Yaschenko, E., Zbicz, K., 2017. Database resources of the national center for biotechnology information. Nucleic Acids Res. 45, D12–D17. https://doi.org/10.1093/nar/gkw1071.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Bateman, A., Martin, M.J., O'Donovan, C., Magrane, M., Alpi, E., Antunes, R., Bely, B., Bingley, M., Bonilla, C., Britto, R., Bursteinas, B., Bye-AJee, H., Cowley, A., Da Silva, A., De Giorgi, M., Dogan, T., Fazzini, F., Castro, L.G., Figueira, L., Garmiri, P., Georghiou, G., Gonzalez, D., Hatton-Ellis, E., Li, W., Liu, W., Lopez, R., Luo, J., Lussi, Y., MacDougall, A., Nightingale, A., Palka, B., Pichler, K., Poggioli, D., Pundir, S., Pureza, L., Qi, G., Rosanoff, S., Saidi, R., Sawford, T., Shypitsyna, A., Speretta, E., Turner, E., Tyagi, N., Volynkin, V., Wardell, T., Warner, K., Watkins, X., Zaru, R., Zellner, H., Xenarios, I., Bougueleret, L., Bridge, A., Poux, S., Redaschi, N., Aimo, L., ArgoudPuy, G., Auchincloss, A., Axelsen, K., Bansal, P., Baratin, D., Blatter, M.C., Boeckmann, B., Bolleman, J., Boutet, E., Breuza, L., Casal-Casas, C., De Castro, E., Coudert, E., Cuche, B., Doche, M., Dornevil, D., Duvaud, S., Estreicher, A., Famiglietti, L., Feuermann, M., Gasteiger, E., Gehant, S., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., Jungo, F., Keller, G., Lara, V., Lemercier, P., Lieberherr, D., Lombardot, T., Martin, X., Masson, P., Morgat, A., Neto, T., Nouspikel, N., Paesano, S., Pedruzzi, I., Pilbout, S., Pozzato, M., Pruess, M., Rivoire, C., Roechert, B., Schneider, M., Sigrist, C., Sonesson, K., Staehli, S., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Veuthey, A.L., Wu, C.H., Arighi, C.N., Arminski, L., Chen, C., Chen, Y., Garavelli, J.S., Huang, H., Laiho, K., McGarvey, P., Natale, D.A., Ross, K., Vinayaka, C.R., Wang, Q., Wang, Y., Yeh, L.S., Zhang, J., 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 45, D158–D169. https://doi.org/10.1093/nar/gkw1099.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242.

Berntsson, R.P.A., Smits, S.H.J., Schmitt, L., Slotboom, D.J., Poolman, B., 2010. A structural classification of substrate-binding proteins. FEBS Lett.. https://doi.org/10.1016/j.febslet.2010.04.043.

Bryksin, A.V., Tomova, A., Godfrey, H.P., Cabello, F.C., 2010. BmpA is a surface-exposed outer-membrane protein of Borrelia burgdorferi. FEMS Microbiol. Lett. 309, 77–83. https://doi.org/10.1111/j.1574-6968.2010.02020.x.

Bryksin, A.V, Godfrey, H.P., Carbonaro, C.A., Wormser, G.P., Aguero-Rosenfeld, M.E., Cabello, F.C., 2005. Borrelia burgdorferi BmpA, BmpB, and BmpD proteins are expressed in human infection and contribute to P39 immunoblot reactivity in patients with Lyme disease. Clin. Diagn. Lab. Immunol. 12, 935–940. https://doi.org/10.1128/CDLI.12.8.935-940.2005.

Buddelmeijer, N., 2015. The molecular mechanism of bacterial lipoprotein modification—How, when and why? FEMS Microbiol. Rev. 39, 246–261. https://doi.org/10.1093/femsre/fuu006.

Caine, J.A., Coburn, J., 2016. Multifunctional and redundant roles of Borrelia burgdorferi outer surface proteins in tissue adhesion, colonization, and complement evasion. Front. Immunol. 7, 1–11. https://doi.org/10.3389/fimmu.2016.00442.

Chung, S.Y., Subbiah, S., 1996. A structural explanation for the twilight zone of protein sequence homology. Structure 4, 1123–1127. https://doi.org/10.1016/S0969-2126(96)00119-0.

Cutler, S.J., 2016. Refugee crisis and re-emergence of forgotten infections in Europe. Clin. Microbiol. Infect. 22, 8–9. https://doi.org/10.1016/j.cmi.2015.10.018.

Cutler, S.J., 2015. Relapsing fever borreliae: a global review. Clin. Lab. Med. 35, 847–865. https://doi.org/https://doi.org/10.1016/j.cll.2015.07.001.

Darcis, G., Hayette, M.P., Bontems, S., Sauvage, A.S., Meuris, C., Van Esbroeck, M., Leonard, P., 2016. Louse-borne relapsing fever in a refugee from Somalia arriving in Belgium. J. Travel Med. 23, 1–3. https://doi.org/10.1093/jtm/taw009.

Davidson, A.L., Shuman, H.A., Nikaido, H., 1992. Mechanism of maltose transport in Escherichia coli: transmembrane signaling by periplasmic binding proteins. Proc. Natl. Acad. Sci. U S A 89, 2360–2364. https://doi.org/10.1073/pnas.89.6.2360.

Deka, R.K., Brautigam, C.A., Biddy, B.A., Liu, W.Z., Norgard, M.V., 2013. Evidence for an ABC-type riboflavin transporter system in pathogenic spirochetes. MBio.. https://doi.org/10.1128/mBio.00615-12.

Deka, R.K., Brautigam, C.A., Yang, X.F., Blevins, J.S., Machius, M., Tomchick, D.R., Norgard, M.V., 2006. The PnrA (Tp0319; TmpC) lipoprotein represents a new family of bacterial purine nucleoside receptor encoded within an ATP-binding cassette (ABC)-like operon in Treponema pallidum. J. Biol. Chem.. https://doi.org/10.1074/jbc.M511405200.

Dobrikova, E.Y., Bugrysheva, J., Cabello, F.C., 2001. Two independent transcriptional units control the complex and simultaneous expression of the Bmp paralogous chromosomal gene family in Borrelia burgdorferi. Mol. Microbiol. 39, 370–378. https://doi.org/10.1046/j.1365-2958.2001.02220.x.

Dowdell, A.S., Murphy, M.D., Azodi, C., Swanson, S.K., Florens, L., Chen, S., Zückert, W.R., 2017. Comprehensive spatial analysis of the Borrelia burgdorferi lipoproteome reveals a compartmentalization bias toward the bacterial surface. J. Bacteriol.. https://doi.org/10.1128/JB.00658-16.

Felsenstein, J., 1985. Confidence limits on phylogenies : an approach using the bootstrap. Author(s): Joseph Felsenstein Published by : Society for the Study of Evolution Stable. Evolution 39, 783–791. http://www.jstor.org/stable/2408678.

Finn, R.D., Attwood, T.K., Babbitt, P.C., Bateman, A., Bork, P., Bridge, A.J., Chang, H.Y., Dosztanyi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G.L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., Mistry, J., Natale, D.A., Necci, M., Nuka, G., Orengo, C.A., Park, Y., Pesseat, S., Piovesan, D., Potter, S.C., Rawlings, N.D., Redaschi, N., Richardson, L., Rivoire, C., Sangrador-Vegas, A., Sigrist, C., Sillitoe, I., Smithers, B., Squizzato, S., Sutton, G., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Wu, C.H., Xenarios, I., Yeh, L.S., Young, S.Y., Mitchell, A.L., 2017. InterPro in 2017—beyond protein family and domain annotations. Nucleic Acids Res. 45, D190–D199. https://doi.org/10.1093/nar/gkw1107.

Fiser, A., 2010. Template-based protein structure modeling. Methods Mol. Biol. 673, 73–94. https://doi.org/10.1007/978-1-60761-842-3_6.

França, T.C.C., 2015. Homology modeling: an important tool for the drug discovery. J. Biomol. Struct. Dyn. 33, 1780–1793. https://doi.org/10.1080/07391102.2014.971429.

Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S., Hanson, M., Van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J., Weidman, J., Utterback, T., Watthey, L., McDonald, L., Artiach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K., Roberts, K., Hatch, B., Smith, H.O., Venter, J.C., 1997. Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. Nature 390, 580–586. https://doi.org/10.1038/37551.

Fraser, C.M., Norris, S.J., Weinstock, G.M., White, O., Sutton, G.G., Dodson, R., Gwinn, M., Hickey, E.K., Clayton, R., Ketchum, K.A., Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J., Khalak, H., Richardson, D., Howell, J.K., Chidambaram, M., Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M.D., Fujii, C., Hatch, B., Roberts, K., Sandusky, M., Weidman, J., Smith, H.O., Venter, J.C., 1998. Complete genome sequence of Treponema pallidum, the syphilis spirochete. Science 281, 375–388 (80-.).

Gorbacheva, V.Y., Godfrey, H.P., Cabello, F.C., 2000. Analysis of the Bmp gene family in Borrelia burgdorferi sensu lato. J. Bacteriol. 182, 2037–2042.

Gupta, R.S., Mahmood, S., Adeolu, M., 2013. A phylogenomic and molecular signature based approach for characterization of the phylum spirochaetes and its major clades: proposal for a taxonomic revision of the phylum. Front. Microbiol. 4, 1–18. https://doi.org/10.3389/fmicb.2013.00217.

Hillis, D.M., Bull, J.J., 2010. In: Society of Systematic Biologists an Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis, 42. Oxford University Press for the Society of Systematic Biologists Stable, pp. 182–192. http://www.jstor.org/stable/2992.

Hovius, J.W.R., van Dam, A.P., Fikrig, E., 2007. Tick-host-pathogen interactions in Lyme borreliosis. Trends Parasitol. 23, 434–438. https://doi.org/10.1016/j.pt.2007.07.001.

Hubálek, Z., 2009. Epidemiology of Lyme Borreliosis. Curr. Probl. Dermatol. 37, 31–50. https://doi.org/10.1159/000213069.

Hytönen, J., Khawaja, T., Grönnroos, J.O., Jalava, A., Meri, S., Oksi, J., 2017. Louse-borne relapsing fever in Finland in two asylum seekers from Somalia. APMIS 125, 59–62. https://doi.org/10.1111/apm.12635.

Illergård, K., Ardell, D.H., Elofsson, A., 2009. Structure is three to ten times more conserved than sequence—A study of structural response in protein cores. Proteins Struct. Funct. Bioinform. 77, 499–508. https://doi.org/10.1002/prot.22458.

Johnson, M.S., Lehtonen, J.V, 2000. Comparision of protein three-dimensional structures. In: Bioinformatics: Sequence, Structure and Databanks, p. 15.

Jones, D.T., 1999. Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. 292, 195–202. https://doi.org/10.1006/jmbi.1999.3091.

Juncker, A., Willenbrock, H., 2003. Prediction of lipoprotein signal peptides in gram negative bacteria. Protein Sci. 12, 1652–1662. https://doi.org/10.1110/ps.0303703.Biosynthesis.

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797. https://doi.org/10.1016/j.jmb.2007.05.022.

Kuehn, B.M., 2013. CDC estimates 3,00,000 US cases of Lyme disease annually. JAMA 310, 1110. https://doi.org/10.1001/jama.2013.278331.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.

Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25, 1307–1320. https://doi.org/10.1093/molbev/msn067.

Lehtonen, J.V, Still, D.-J., Rantanen, V.-V., Ekholm, J., Björklund, D., Iftikhar, Z., Huhtala, M., Repo, S., Jussila, A., Jaakkola, J., Pentikäinen, O., Nyrönen, T., Salminen, T., Gyllenberg, M., Johnson, M.S., 2004. BODIL: a molecular modeling environment for structure-function analysis and drug design. J. Comput. Aided Mol. Des. 18, 401–419.

Lescot, M., Audic, S., Robert, C., Nguyen, T.T., Blanc, G., Cutler, S.J., Wincker, P., Couloux, A., Claverie, J.M., Raoult, D., Drancourt, M., 2008. The genome of Borrelia recurrentis, the agent of deadly louse-borne relapsing fever, is a degraded subset of tick-borne Borrelia duttonii. PLoS Genet. 4. https://doi.org/10.1371/journal.pgen.1000185.

Liang, F.T., Nelson, F.K., Fikrig, E., 2002. Molecular adaptation of Borrelia burgdorferi in the murine host. J. Exp. Med. 196, 275–280.

Liu, C.E., Liu, P.Q., Wolf, A., Lin, E., Ames, G.F.L., 1999. Both lobes of the soluble receptor of the periplasmic histidine permease, an ABC transporter (traffic ATPase), interact with the membrane- bound complex. J. Biol. Chem. 274, 739–747. https://doi.org/10.1074/jbc.274.2.739.

Maghrabi, A.H.A., Mcguffin, L.J., 2017. ModFOLD6: an accurate web server for the global and local quality estimation of 3D protein models. Nucleic Acids Res. 45. https://doi.org/10.1093/nar/gkx332.

Mao, B., Pear, M.R., McCammon, J.A., Quiocho, F.A., 1982. Hinge-bending in L-arabinose-binding protein. J. Biol. Chem. 257, 1131–1134.

Maqbool, A., Horler, R.S.P., Muller, A., Wilkinson, A.J., Wilson, K.S., Thomas, G.H., 2015. The substrate-binding protein in bacterial ABC transporters: dissecting roles in the evolution of substrate specificity. Biochem. Society Trans. 43, 1011–1017. https://doi.org/10.1042/BST20150135.

Nordstrand, A., Bunikis, I., Larsson, C., Tsogbe, K., Schwan, T.G., Nilsson, M., Bergström, S., 2007. Tickborne relapsing fever diagnosis obscured by Malaria, Togo. Emerg. Infect. Dis. 13, 117–123. https://doi.org/10.3201/eid1301.060670.

Pal, U., Wang, P., Bao, F., Yang, X., Samanta, S., Schoen, R., Wormser, G.P., Schwartz, I., Fikrig, E., 2008. Borrelia burgdorferi basic membrane proteins A and B participate in the genesis of Lyme arthritis. J. Exp. Med. 205, 133–141. https://doi.org/10.1084/jem.20070962.

Pearson, W.R., 2013. An introduction to sequence similarity ("homology") searching. Curr. Protoc. Bioinform. 1–8. https://doi.org/10.1002/0471250953.bi0301s42.

Pettersson, J., Schrumpf, M.E., Raffel, S.J., Porcella, S.F., Guyard, C., Lawrence, K., Gherardini, F.C., Schwan, T.G., 2007. Purine salvage pathways among Borrelia species. Infect. Immun. 75, 3877–3884. https://doi.org/10.1128/IAI.00199-07.

Radolf, J.D., Caimano, M.J., Stevenson, B., Hu, L.T., 2012. Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes. Nat. Rev. Microbiol. 10, 87–99. https://doi.org/10.1038/nrmicro2714.

Radolf, J.D., Deka, R.K., Anand, A., Šmajs, D., Norgard, M.V., Yang, X.F., 2016. Treponema pallidum, the syphilis spirochete: making a living as a stealth pathogen. Nat. Rev. Microbiol. 14, 744–759. https://doi.org/10.1038/nrmicro.2016.141.

Ramamoorthy, R., McClain, N.A., Gautam, A., Scholl-Meeker, D., 2005. Expression of the bmpB gene of Borrelia burgdorferi is modulated by two distinct transcription termination events. J. Bacteriol. 187, 2592–2600. https://doi.org/10.1128/JB.187.8.2592-2600.2005.

Ramamoorthy, R., Povinelli, L., Philipp, M.T., 1996. Molecular characterization, genomic arrangement, and expression of bmpD, a new member of the Bmp class of genes encoding membrane proteins of Borrelia burgdorferi. Infect. Immun. 64, 1259–1264.

Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res.. 42 https://doi.org/10.1093/nar/gku316.

Rowley, J., Toskin, I., Ndowa, F., 2012. Global incidence and prevalence of selected curable sexually transmitted infections – 2008.

Salentin, S., Schreiber, S., Haupt, V.J., Adasme, M.F., Schroeder, M., 2015. PLIP: fully automated protein – ligand interaction profiler. Nucleic Acids Res. 43, W443–W447. https://doi.org/10.1093/nar/gkv315.

Šali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815. https://doi.org/10.1006/jmbi.1993.1626.

Scheepers, G.H., Lycklama A Nijeholt, J.A., Poolman, B., 2016. An updated structural classification of substrate-binding proteins. FEBS Lett. 590, 4393–4401. https://doi.org/10.1002/1873-3468.12445.

Simpson, W.J., Burgdorfer, W., Schrumpf, M.E., Karstens, R.H., Schwan, T.G., 1991. Antibody to a 39-kilodalton Borrelia burgdorferi antigen (P39) as a marker for infection in experimentally and naturally inoculated animals. J. Clin. Microbiol. 29, 236–243.

Simpson, W.J., Cieplak, W., Schrumpf, M.E., Barbour, A.G., Schwan, T.G., 1994. Nucleotide sequence and analysis of in Borrelia burgdorferi encoding the P39 antigen the gene immunogenic. FEMS Microbiol. Lett. 119, 381–388. https://doi.org/10.1111/j.1574-6968.1994.tb06917.x.

Simpson, W.J., Schrumpf, M.E., Schwan, T.G., 1990. Reactivity of human Lyme borreliosis sera with a 39-kilodalton antigen specific to Borrelia burgdorferi. J. Clin. Microbiol. 28, 1329–1337.

Sippl, M.J., 1993. Recognition of errors in three dimensional structures of proteins. Proteins Struct. Funct. Bioinform.. https://doi.org/10.1002/prot.340170404.

Soltis, D.E., Soltis, P.S., 2014. Update on Molecular Systematics The Role of Phylogenetics in Comparative Genetics 1 132, 1790–1800. https://doi.org/10.1104/pp.103.022509.and

Soltis, D.E., Soltis, P.S., 2003. Applying the bootstrap in phylogeny reconstruction. Stat. Sci. 18, 256–267. https://doi.org/10.1214/ss/1063994980.

Sykes, R.A., Makiello, P., 2014. An estimate of Lyme borreliosis incidence in Western Europe. Res Medica 22, 76. https://doi.org/10.2218/resmedica.v22i1.743.

Verma, A., Brissette, C.A., Bowman, A., Stevenson, B., 2009. Borrelia burgdorferi bmpA is a laminin-binding protein. Infect. Immun. 77, 4940–4946. https://doi.org/10.1128/IAI.01420-08.

Wallner, B., Elofsson, A., 2003. Can correct protein models be identified? Protein Sci. 12, 1073–1086. https://doi.org/10.1110/ps.0236803.

Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35, 407–410. https://doi.org/10.1093/nar/gkm290.

Yang, X., Izadi, H., Coleman, A.S., Wang, P., Ma, Y., Fikrig, E., Anguita, J., Pal, U., 2008. Borrelia burgdorferi lipoprotein bmpA activates pro-inflammatory responses in human synovial cells through a protein moiety. Microbes Infect. 10, 1300–1308. https://doi.org/10.1016/j.micinf.2008.07.029.

Zhao, H., Liu, A., Cui, Y., Liang, Z., Li, B., Bao, F., 2017. Borrelia burgdorferi basic membrane protein A could induce chemokine production in murine microglia cell line BV2. Microb. Pathog. 111, 174–181. https://doi.org/10.1016/j.micpath.2017.08.036.

Zückert, W.R., 2014. Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. Biochim. Biophys. Acta 1843, 1509–1516. https://doi.org/10.1016/j.bbamcr.2014.04.022.