Characterisation and structural studies of a superoxide dismutase and OmpA-like proteins from *Borrelia burgdorferi sensu lato*

Gemma Brown BSc

A thesis submitted to the University of Huddersfield

in partial fulfilment of the requirements for

the degree of Doctor of Philosophy

July 2016

Copyright declaration

For my family

"Wine is the most healthful and most hygienic

of beverages."

Louis Pasteur, 1822 – 1895.

Abstract

Lyme borreliosis is the most common tick-borne, human infection across the Northern hemisphere. The agent responsible, *Borrelia burgdorferi sensu lato (s.l.)* covers a family of *Spirochaetes* with unique characteristics which are shared by both Gram-negative and Gram-positive bacteria. The outer membrane (OM) is rich in lipoproteins but contains a relatively low density of integral membrane proteins (OMPs), of these OMPs very few have been identified and even fewer are well characterised. The OmpA-like transmembrane domain defined by the Pfam family PF01389 is a 8-stranded membrane spanning β-barrel and is well conserved among Gram-negative bacteria but to date remains unknown in *Spirochaetes*. Building from previous computational work which had sought to identify possible OMPs from *B. burgdorferi s.l.* four OmpA-like proteins, BAPKO_0422 (*Borrelia afzelii*), BB_0562, BB_0406 (*B. burgdorferi*) and BG0408 (*Borrelia garinii*) have been identified and structurally characterised. The four proteins are encoded by chromosomal genes and highly conserved between *Borrelia* species and may be of diagnostic or therapeutic value. Structural characterisation by both circular dichroism and small angle X-ray scattering suggests these four proteins adopt a compact globular structure rich in β-strand (~40%) with *Ab initio* molecular envelopes resembling a cylindrical peanut shape with dimensions of ~25x45 Å consistent with an 8-stranded β-barrel. The present work demonstrates that BAPKO_0422 can bind human factor H (hfH) and some evidence for a further interaction between the BAPKO_0422 protein and heparin. The interaction with hfH may contribute to the spirochaete's immune evasion mechanisms by the inhibition of the complement response.

The zoonotic life-cycle of *Borrelia* and challenges by the host's immune system causes an ever changing environment which often leads to fluctuations of O2 exposure. Although *B. burgdorferi s.l.* have a distinct lack of metabolic systems including peroxidases and catalase enzymes the *Spirochaetes* genome does encode a single superoxide dismutase gene (*sodA - bb_0157*). Previously assigned as a Fe-SOD there has been some debate whether this protein requires iron or manganese as a co-factor. The present work demonstrates that the *B.* burgdorferi enzyme SodA requires manganese for activity and does not display cambialistic behaviour. Structural and proteomic characterisation suggests the *B. burgdorferi* SodA enzyme shares significant sequence similarity to a superoxide dismutase from *Thermus thermophilus*.

Table of Contents

List of Figures

List of Tables

List of Equations

Abbreviations

ACA	Acrodermatitis chronica atrophicans
AP	Alternative Pathway
AV	Atrioventricular
BBB	Blood-brain barrier
BbCRASP	*B. burgdorferi s.l.* Complement Regulator Acquiring Surface Protein
BLAST	Basic Local Alignment Search Tool
BMEC	Brain microvascular endothelial cell
C_8E_4	n-octyltetraoxyethylene
CD	Circular Dichroism
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CM	Cytoplasmic membrane
CNS	Central nervous system
CP	Classical Pathway
Dbp	Decorin binding protein
DDM	n-Dodecyl β-D-maltoside
Erp	OspE Related Protein
fH	Factor H
fHR-1	Factor H-related protein
Fn	Fibronectin
FPLC	Fast Protein Liquid Chromatography
HMM	Hidden-Markov Model
IMAC	Immobilised Metal Affinity Chromatography
LA	Lyme arthritis
LBRF	Louse-borne relapsing fever
LCP	Lipid Cubic Phase
LDAO/DDAO	N,N-Dimethyldodecylamine *N*-oxide
LNB	Lyme neuroborreliosis
LPS	Lipopolysaccharide
LS	Light Scattering
MBL	Mannose-binding lectin
MS	Mass Spectrometry
MSA	Multiple Sequence Alignment
MO	Monoolein

NMR	Nuclear Magnetic Resonance Spectroscopy
OG	n-octyl-β-D-glucoside
OM	Outer membrane
OMP	Outer membrane protein
Osp	Outer surface protein
PLDS	Post-treatment Lyme disease syndrome
*R*g	Radius of Gyration
RI	Refractive Index
SAXS	Small-Angle X-ray Scattering
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
SEC	Size Exclusion Chromatography
SEC-MALLS	Size Exclusion Chromatography-Multi Angle Laser Light Scattering
s.s	*Sensu stricto*
s.l.	*Sensu lato*
TBRF	Tick-borne relapsing fever

Acknowledgements

First and foremost I thank my supervisor, Dr Richard Bingham for providing me not only the opportunity to pursue this work, but also continuous support and advice throughout my PhD which will be sadly missed.

I thank the University of Huddersfield for providing a fee waiver and scholarship and a place to undertake this research. I also thank the biological sciences technical staff who have provided me with lots of assistance over the past few years. Many thanks also to Dr Andrew Leech (University of York) for assistance with circular dichroism, Dr Gabi Margos (University of Bath) and Dr Volke Fingerle for providing *Borrelia* DNA and to Diamond Light Source for funding and supporting SAXS experiments (Session ID: SM12308-1 on B21 28/05/2015).

I thank all my family for supporting me both emotionally and financially during my years in education, for what now feels like forever, but lastly my heartfelt thanks to Martin, without whose support this thesis would have been written in half the time but with half the happiness.

Chapter 1: Introduction

1.1 Lyme Disease

1.1.1 A brief history of Lyme disease

Lyme borreliosis or Lyme disease is the most common arthropod-borne, human infection reported throughout the northern hemisphere (CDC, 2013). Since the beginning of the twentieth century, symptoms associated with the disorder were well documented; however the etiological agent remained unidentified. During the 1970's, unusual clusters of oligoarthritis, mainly identified in juveniles from rural communities were reported around the town of Lyme, situated in southeastern Connecticut, USA (Steere et al., 1977). The annual epidemics which peaked during the summer months from as early as 1972, ran somewhat undetected and were largely misdiagnosed as juvenile rheumatoid arthritis (Steere et al., 1972). Following the assistance from nearby Yale University, which was initially requested by a group of shrewd parents, the observation that a number of patients had presented with a distinctive, expanding skin lesion, provided the key in identifying the agent responsible for what is now described as a global health concern. The now familiar bulls eye rash or erythema migrans had previously been described in Northern Europe, following a bite from the sheep tick, *Ixodes scapularis,* which was already believed to be caused by an unknown infectious agent (Hellerström, 1930). This link propelled an intensive investigation, leading to the isolation of a novel spirochaete from the North American deer tick *Ixodes scapularis,* so named *Borrelia burgdorferi* after its co-discoverer Willy Burgdorfer (Burgdorfer et al., 1982).

From the initial isolation and identification of the spirochaete, no less than 20 closely related species were identified with this number continuing to grow and have become known as *B.burgdorferi sensu lato (s.l.)*. Within *B. burgdorferi s.l.* at least three genospecies are pathogenic to humans, *B. afzelii* and *B. garinii* in Europe and *B.burgdorferi sensu stricto,* predominantly within the United States however the number of pathogenic strains is growing.

Borrelia are highly specialised pathogens with the ability to tightly regulate gene expression in response to their environment, allowing them to achieve a complex lifecycle between invertebrate vectors and vertebrate hosts (Fraser et al., 1997). The transmission of *B.burgdorferi* requires the efficient migration, from the tick's gut to the salivary glands, during the invertebrate's blood meal, which stages the penetration of the pathogen into the vertebrate's tissues (Ribeiro et al., 1987). Subsequent colonisation of uninfected ticks from the host

completes the complex enzootic cycle and highlights the pathogens requirement to significantly adapt to the often changing environment (Fisher et al., 2005).

Over the past few years Lyme disease has become a hot topic, especially within the US, with the 2012 presidential candidate Mitt Romney including the disease within his campaign strategy (The New Yorker, 2012). The two medical societies, The Infectious Diseases Society of America (IDSA) and The International Lyme and Associated Diseases Society (ILADS) are at arms following the development of conflicting clinical and patient recommendations. This conflict has rapidly escalated with the IDSA accusing ILADS as being involved in an 'axis of evil' by their association and recommendation of misguided physicians and controversial treatment strategies. In retaliation ILADS believe that the IDSA are ignoring crucial scientific information and that they have conflicts of interest with commercial value when publishing guidelines. The pair appear to have formed a stance with the IDSA against and the ILADS for the belief of chronic infections which require long antibiotic treatment periods with both sides publishing research in support and against pro-longed therapies. Some researchers have attempted to bring clarity to their work in the form of independent review articles the most recent by Borchers, 2014 and Sanchez, 2015. Patients unfortunately are stuck in the middle and are often attracted to treatments which sound promising but have little scientific significance when current diagnostic testing and standard treatments fail.

1.1.2 Symptoms

Early symptoms of localised infection are often seen within three to 30 days from transmission and often take the form of erythema migrans (EM), more commonly known as the 'bull's eye rash' as shown in figure 1.1 (Murray and Shapiro, 2010, Tibbles et al., 2007, DePietropaolo et al., 2006). Up to 80 per cent of patients infected, present with the characteristic rash, which is commonly defined as a circular or oval-shaped skin lesion, which may expand outwards and range in size between 5 to 70cm (Tibbles et al., 2007). Adults most commonly present with EM on the legs, buttocks or groin area and less frequently on the arms and trunk. Presentations of EM upon the head are mostly seen in children at a rate of 20-40% of cases (Berglund et al., 1995 and Altpeter et al., 2013). Interestingly the typical bull's eye effect with a central clearing is only seen in 19% of American cases compared to 80% of European patients (Tibbles and Edlow, 2007).

EM was previously considered exclusive to Lyme disease. However the USA has recorded patients presenting with EM who have been diagnosed with an illness similar to that of Lyme borreliosis known as southern tick-associated rash illness (STARI) which is transmitted by the lone star tick (*A. americanum*) by an unknown agent (Tibbles and Edlow, 2007). Fortunately for most patients presenting with an EM rash antibiotic therapy is often successful and further complications are rare. However, for those patients where an EM is not present, not seen or they are misdiagnosed, disseminated infection can lead to dangerous complications (Borchers et al., 2014). A full summary of symptoms and their associated disease stages can be reviewed in table 1.1.

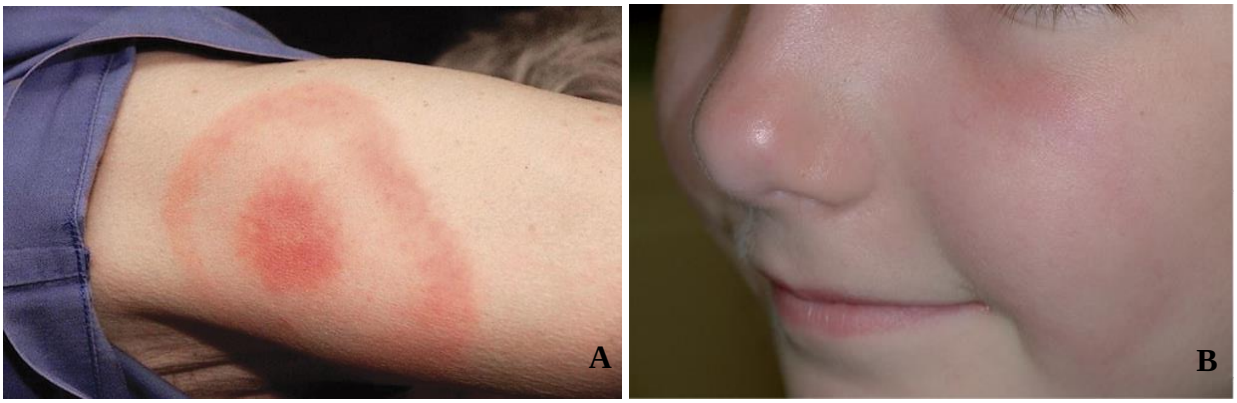


Figure 1.1 Examples of the early manifestation of Lyme disease, Erythema migrans. **A**. The classical bull's eye rash on the arm of an adult patient (Ogden et al., 2008). **B**. A faint annular EM on a paediatric patient (Glatz et al., 2014).

Presentation of Lyme borreliosis following infection can be clearly grouped into three specific stages – early localised infection, early disseminated infection and late persistent infection, also described as primary, secondary and tertiary (Borchers et al., 2014). There is also a further stage which has been coined as post-Lyme syndrome, whether this is indeed still an active infection or a resulting immune disorder is still up for debate (Wormser et al., 2006, Murray and Shapiro, 2010, Borchers et al., 2014).

Following infection, systemic symptoms may become apparent. These include vague symptoms such as headache, fatigue, a stiff neck, fever, nausea, myalgias and dysesthesia (Tibbles and Edlow, 2007). A more obvious symptom is the presence of a further EM lesion

away from the original site of transmission. *Borrelial* lymphocytoma, lymphocytoma cutis or lymphadenosis benigna cutis is observed in around 4.5% of European cases of early disseminated Lyme borreliosis and most often in children (Altpeter et al., 2013, Berglund et al., 1995, Huppertz et al., 1999, Christova and Komitova, 2004, Dhôte et al., 2001, Lipsker et al., 2001). It appears as a bluish-red nodular swelling (figure 1.2) and is often seen on the earlobe of infected children or upon the nipple of an infected adult and can be painful to the touch (Maraspin et al., 2002, Strle at al., 1992, Colli et al., 2004). It can be seen from weeks to months following infection and sometimes during the EM period. Although lymphocytoma is not exclusive to *Borrelia* infection the frequency of its presentation is highest in Lyme endemic areas (Borchers et al., 2014).

Figure 1.2. An example of early disseminated Lyme disease, *Borrelia* lymphocytoma. An example of *Borrrelia* lymphocytoma on the ear of an Austrian paediatric patient. (Glatz et al., 2014)

When EM is not present or is no longer present, Lyme neuroborreliosis (LNB) is often the presenting complaint of disseminated infection and can develop anytime between days to 3 months following exposure (Hubálek, 2009). Studies suggest that 10-15% of patients who are not treated for EM will develop neurological symptoms (Steere et al., 1978, Asbrink et al., 1986) with the highest rates seen in children and within the over 50 age group (Berglund et al., 1995, Huppertz et al., 1999, Hansen & Lebech., 1992, Henningsson et al., 2010). Presentation of neuroborreliosis appears to have a male to female ratio of 1.5:1 in Europe (Hansen & Lebech., 1992, Oschmann et al., 1998, Bremell & Hagberg., 2011, Henningsson et al., 2010

and Beglund et al., 2002). Neurological manifestations vary but most common examples are, cranial neuropathies such as facial nerve palsy, lymphocytic meningitis, cranial neuritis, radiculoneurtitis and encephalitis (Greco et al., 2012, Reik et al., 1979, Dhôte et al., 2000, Hansen & Lebech., 1992, Oschmann et al., 1998, Henningsson et al., 2010). A breakdown of American cases reported to the CDC showed 9% of LNB patients presented with Bell's palsy, 4% showed radiculoneuritis, 1% developed meningitis and a further 1% encephalitis (CDC, 2013). European datum suggests that 16-23% of infected patients present with some form of LNB. Of these 70-90% of these adult patients complained of severe radiculitic pain often at its peak during the night (Dhôte et al., 2000, Hansen & Lebech., 1992 and Oschmann et al., 1998) and cranial neuritis was seen to affect 40-50% of adult patients mainly as Bell's palsy (Dhôte et al., 2000, Hansen & Lebech., 1992, Oschmann et al., 1998, Henningsson et al., 2010). CSF lymphocytic pleocytosis is also common in European patients with neuroborreliosis (Dhôte et al., 2000, Hansen & Lebech., 1992, Oschmann et al., 1998, Henningsson et al., 2010) and has become part of the diagnostic criteria for LNB. Other rare but observed conditions seen in European cases are, progressive encephalomyelitis, chronic meningoencephalitis, meningomyelitis and radiculomyelitis (Hansen & Lebech., 1992, Oschmann et al., 1998).

Although cardiac involvement is rare, Lyme cardioborreliosis is observed in upto 1% of both American (CDC, 2013) and European cases reported, often associated with early disseminated disease. The condition can become apparent anywhere between a few days to three months following infection (Berglund et al., 1995, Lesnyak et al., 1998, Huppertz et al., 1999 and Christova et al., 2004). Statistically Lyme carditis is mostly seen in the 20 to 40 years age group with a ratio of 3:1, male to female. It presents most commonly as fluctuating disturbances in atrioventricular (AV) conduction (Steere et al., 1980, van der Linde., 1991 and Costello et al., 2009) and the block is most commonly above the bundle of His, within the AV node (Steere et al., 1980 and van der Linde., 1991) although other areas within the AV conduction system can be affected. Based on recorded cases, half resulted in a total heart block with one third of these patients requiring a temporary pacemaker (Costello et al., 2009, McAlister et al., 1989 and Midttun et al., 1997). Other reported cardiac manifestations include, endocarditis, pericarditis, myocarditis, pericardial effusion and congestive heart failure (Berglund et al., 1995, Lesnyak et al., 1998, Huppertz et al., 1999, Borchers et al., 2014).

Clinical manifestations of Lyme borreliosis		
Stage	**Other names**	**Manifestations**
Stage 1	Early localised	Erythema migrans (EM) Flu-like symptoms, headache, myalgia, arthralgia, local lymphadenopathy. *Borrelia* lymphocytoma – commonly of the earlobe or nipple
Stage 2	Early disseminated	Multiple erythema migrans Sweats, myalgia, arthralgia Viral-like meningitis or encephalitis Cranial nerve palsy Mononeuritis multiplex Radiculopathies Cardiac conduction disturbances Arthritis Myocarditis Pericarditis Cardiomyopathy and heart failure
Stage 3	Late disseminated	Oligoarthritis Acrodermatitis chronica atrophicans
Stage 4	Post Lyme syndrome or Chronic Lyme disease	Fatigue Joint and muscle pain Sleep issues and depression Auto-immune responses Neuropathy

Table 1.1. Stages and clinical manifestations of Lyme borreliosis. Table of disease progression with related symptoms. The table has been adapted from the British Journal of Hospital Medicine with an additional stage representing chronic symptoms often associated with Post Lyme Disease Syndrome or chronic Lyme disease.

Later cutaneous manifestations of Lyme borreliosis are predominantly observed in women from middle age upwards and take form as acrodermatitis chronica atrophicans or ACA. This predisposition towards females remains unexplained but the condition appears as a bluish-red

discolouration of the skin normally on the upper side of the feet or hands (Asbrink et al., 1986). It has been reported that this condition often occurs on the same area in which the original EM was observed years earlier although only 20% of these patients gave a history of the original EM manifestation (Asbrink et al., 1986). ACA may spread to other extremities of the body and become symmetrical and is rarely seen to resolve quickly with symptoms often remaining for several years followed by atrophy of the skin (Asbrink & Hovmark., 1988). Development of a sensory polyneuropathy is also seen in 30-40% of patients following ACA (Kindstrand et al., 1997, Asbrink & Hovmark., 1988 and Kristoferitscho et al., 1988). Development of ACA is fairly rare and occurs in 1-7% of European Lyme disease patients (Berglund et al., 1995, Altpeter et al., 2013 and Huppertz et al., 1999). However has not yet been reported in Chinese (Ai et al., 1988) or Russian cases (Lesnyak et al., 1998) and it is also particularly rare within the USA.

Following infection, 60% of patients who present with EM but who do not receive appropriate treatment develop recurrent episodes of Lyme arthritis (LA) (Steere et al., 1987). However this condition is extremely rare in patients who do receive a suitable antibiotic therapy (Borchers et al., 2014). This presents problems for those patients who don't recognise an EM or who do not present with the typical early symptoms that would be easily recognised by a primary care physician. Arthritic conditions can develop during the early infection stage anytime from days to a few months after infection but is also seen as a repeating late manifestation which can occur years later. Typical LA is described as brief but reoccurring monoarticular or asymmetric oligoarticular swelling and pain of the large joints or migratory polyarthritis of both the small and large joints (Steere et al., 1987, Huppertz et al., 1999, Gerber et al., 1998). The condition often resolves even in untreated patients however some cases have shown evidence of chronic synovitis and permanent joint destruction (Steere et al., 1987).

A flurry of recent studies has suggested a possible link or correlation with LNB and the development of Alzheimer's disease (AD) (Bu et al., 2014, Mawanda and Wallace, 2013, Halperin, 2011 and Miklossy, 2008). Past research has elucidated that infectious burden plays a role in cardiovascular disease and stroke with some evidence emerging that the overall burden of prior infection, including an infection with *Borrelia* increases the likelihood of developing AD (Bu et al., 2014). This topic still remains controversial with much further investigation required.

1.1.3. Diagnosis

Diagnosis of Lyme borreliosis remains discordant. For patients that present with EM, alongside a history of potential tick exposure, diagnosis is relatively straightforward. However the absence of EM, for example during the later disease stages, renders the diagnosis reliant upon laboratory confirmation of infection. Genetic studies following the completion of *B.burgdorferi's* genome (Fraser et al., 1997) suggested an unusual absence of biosynthetic pathways (discussed in section 1.3.4 Genomic Complexity), highlighting the pathogens requirement to obtain nutrients from its environment (Fraser et al., 1997). This distinct lack of metabolic systems, makes the spirochete difficult to culture under normal conditions and *in vitro* growth is only obtained through the use of highly enriched culture media (Barbour, 1984).

Both direct and indirect methods have been used within the laboratory setting to assist diagnosis of Lyme borreliosis, direct methods include culture or detection of the pathogen specific proteins and indirect diagnosis involves detection of antibodies raised against the invading bacterium (Murray et al., 2010). The gold standard is the isolation of *Borrelia spp*. by culture followed by PCR-based approaches to confirm the *Borrelia* strain. In many cases culture is not carried out as the turn-around time is often 2-6 weeks often after a clinical treatment decision is made and the test often returns negative with an overall sensitivity of 40-70% for a sample taken from an EM (Strle et al., 1996 & Smith et al., 2002) and a 3-17% sensitivity for CSF samples (Strle et al., 2006 and van Dam et al., 1993). This low sensitivity can be attributed to the difficulty in culturing the bacteria or the relatively low levels of *Borrelia* within the fluid taken.

PCR based techniques based upon the principle of recognising *Borrelia* DNA can be useful especially if this DNA is found within synovium samples from patients presenting with Lyme Borreliosis. However this isn't always an indication of active disease and again suffers from the same problems with a sensitivity of 75-80% from EM samples (Liveris et al., 2002 and O'Rourke et al 2013), 15-30% from the CSF (Lebech et al., 2002 and Nocton et al., 1996) and 60-85% from synovial fluid samples (Nocton et al., 1994 and Jones et al., 2009).

A single serological test can often yield a false positive result as many individuals are infected with spirochaete bacteria prior to exposure to *Borrelia* and spirochaetal oral infection is common. Within the UK a two-tiered testing method is recommended. This diagnostic method uses an enzyme linked immunosorbent assay (ELISA) followed by a confirmatory western

blot. The initial ELISA is a C6 antigen based assay which is combined for both IgG and IgM followed by a confirmatory western blot for IgG and IgM separately. In order for a patient to be diagnosed with Lyme disease in absence of EM, both of the tests must be returned positive alongside Lyme disease symptoms. Again sensitivity is problematic as outlined in table 1.2, and many people presenting with EM in Europe never seroconvert. Antibodies against *Borrelia* are often slow to develop with IgM undetectable until after 1-2 weeks following infection and IgG often undetectable until 6 weeks has passed (Berglund et al., 1995, Aguero-Rosenfeld et al., 1993, Engstrom et al., 1995, Steere et al., 2008, Glatz et al., 2008, Aguero-Rosenfeld et al., 1996 and Craft et al., 1984). Following the eventual generation, IgM and IgG antibodies may persist for many years (Engstrom et al., 1995 and Aguero-Rosenfeld et al., 1996) making it impossible to discriminate between previous and new cases of infection. Interestingly recent research has identified that the rate of seropositivity is connected to the duration of symptoms prior to seeking diagnosis and treatment (Berglund et al., 1995, Aguero-Rosenfeld et al., 1996, Tveitnes et al., 2009, Cerar et al., 2010). This may suggest that early antibiotic treatment plays some role in preventing development of seropositivity particularly in Europe (Glatz et al.,2006 and Stanek and Kahl, 1999).

Test Type	**% Reactivity for patients with:**			
	EM Acute phase	EM Convalescent phase	Neurological presentations	Arthritic presentations
Whole-cell ELISA	33-49	76-86	79	100
IgM Western blot	43-44	75-84	80	16
IgG Western blot	0-13	15-21	64-72	96-100
Two-tier test method	29-40	29-78	87	97

Table 1.2 – Diagnostic reactivities of *B. burgdorferi* antibody detection within US patients using various methods. Table adapted from a review by Aguero-Rosenfeld et al., 2005. The table demonstrates the huge range in reactivities observed for the current *B. burgdorferi* testing methods during different stages of clinical presentation and highlights the need to improve the current methods used.

1.1.4. Treatment and vaccination

Treatment of Lyme borreliosis varies throughout affected countries, once diagnosed correctly, antibiotic therapy is often employed and can prevent development of other complications (Borchers et al., 2014). Effective antibiotics include β-lactams, tetracyclines and in some cases macrolides are useful (Borchers et al., 2014). The best procedure still remains unclear which leads to variation of treatment strategies across the globe. It is becoming more evident that not all individuals respond the same and not all *Borrelia* species react equally to the current antibiotics in use (Preac Mursic et al., 1996). Treatment guidance has been proposed from several organisations around the world and is mostly based upon available evidence from randomised controlled trials (RCTs). On the whole, recommendations are similar with the major differences lying within dosage and treatment lengths with ILADS advocating harsher and more aggressive treatment strategies. However, the justification for these extended treatments are based on relatively old studies possibly giving an unbalanced opinion by ignoring much newer research which often does not support the use of prolonged treatment strategies (Borchers et al., 2014).

As the risk of acquiring *Borrelia* infection following a tick bite remains low even when the tick itself is infected, prophylaxis is not currently recommended under normal circumstances. However, a US randomised placebo-controlled trial of prophylaxis by means of a single dose of 200 mg of doxycycline within 72 hours of tick removal was effective at preventing the development of initial stage symptoms (Nadelman et al., 2001). Alike the American study a Russian trial found they could reduce the infection rate by a factor of 11 when administering 200 mg of doxycycline for upto 5 days following a tick bite (Korenberg et al., 1996).

The European National Neurological Society (EFNS) recommends an oral course of doxycycline or intravenous administration of ceftriaxone (Steere et al., 2004) over a period of 2-4 weeks, for adult patients presenting with early dissemination stages of infection (Rosa et al., 2005). Some studies have demonstrated the requirement of multiple, long term courses of antibiotics for those patients diagnosed with chronic Lyme. However, this is not always effective once the spirochaete has gained access to the central nervous system (Livengood et al., 2006). Such extended antibiotic treatment is also considered unsuitable or ineffective for patients presenting with symptoms which have persisted for over six months following infection (Mygland et al., 2009). This condition is often described as post Lyme disease syndrome. Compilation of several large RCTs have not shown any evidence that the extension

of doxycycline treatment from 10 to 20 days improves average patients outcomes (Wormser et al., 2003, Stupica et al., 2012 and Kowalski et al., 2010). Many of these studies have noted that in a small number of cases of extended treatment some patients did benefit. However, such benefits need to outweigh the risks of aggressive therapy. Arthritic symptoms that persist for over three months following antibiotic treatment are often referred to as treatment resistant LA and are best tackled using a different strategy other than antibiotics especially when synovial fluid is not positive for *borrelial* DNA. These include, anti-inflammatory drugs, intra-articular corticosteroids or disease modifying anti-rheumatic drugs (Wormser et al., 2006).

A commercially available vaccine was developed in 1998 (Steere et al., 1998). However, it was withdrawn three years later due to decreasing sales following controversy over possible side effects (Krupka et., 2011). The outer surface protein A (OspA) vaccine was administered in three 30µg doses, following the final dose 99% of the participants tested positive for the OspA antibody. The study proposed the vaccine was 76% effective against *B. burgdorferi sensu lato* (Steere et al., 1998, Nigrovic et al., 2006). Further production of alternative vaccines has not yet been successful, primarily due to antigenic variation of outer surface proteins between different strains (Krupka et al., 2011).

1.1.5. Post-Lyme disease syndrome (PLDS)

Following antibiotic therapy, 3-27% of patients who presented with EM complain of remaining symptoms (Borchers et al., 2014). This proportion increases to 10-40% for those who were diagnosed with LNB (Borchers et al., 2014). Residual symptoms often include, headache, fatigue, irritability, arthralgias, myalgias, stiff neck, paresthesias (Berglund et al., 2002, Ljøstad & Mygland et al., 2009, Borg et al., 2005, Eikeland et al., 2012, Eikeland et al., 2011, Benke et al., 1995 and Vrethem et al., 2002) and issues with concentration and memory (Shadick et al., 1994, Ravdin et al., 1996 and Klempner et al., 2001). Neuropsychological studies of previous adult LNB patients up to 30 months following the resolution of infection have shown reduced verbal and visual memory, deficits in attention/executive function and reduced processing speed when compared to healthy controls (Eikeland et al., 2012 and Benke et al., 1995) however this is not seen in paediatric patients.

As discussed in a recent review article (Borchers et al., 2014) several studies have stated that further antibiotic therapy for these residual symptoms is ineffective (Mygland et al., 2009), not

useful (Wormser et al., 2006) or has not shown sustained benefit (British Infection Association, 2011). The review makes an interesting point that all three of these statements are false as in two of the same studies a subset of the patients did indeed show a significant improvement which was sustained. The review calls into question our lack of understanding of the human microbiome in Lyme patients and suggests this may influence the symptom of PLDS. Evidence for PLDS remains unclear and the term seems to now describe any symptoms a patient may suffer following an adequate antibiotic treatment. There is no current standardised case definition for the syndrome and much further study is required in order to identify patients with active recurring borreliosis from those who are not suffering from an active infection but may be displaying post Lyme symptoms such as autoimmune disorders or after effects following neuroborreliosis.

1.1.6. Incidence

Lyme borreliosis or Lyme disease is the most common arthropod-borne, human infection reported throughout the northern hemisphere and is becoming increasingly more common. During 2001 within the UK there were a total of 268 cases of laboratory confirmed Lyme borreliosis, giving a mean annual rate of infection of 0.50 per 100,000 people (HPA, 2013). By 2011 using the latest available HPA statistics, this number had increased to 959 or 1.73 per 100,000 (HPA. 2013). Since 2001 the number of laboratory confirmed cases has steadily increased and this increase can be attributed to several factors. Proposed factors vary and include, changes in the population size of *I. ricinus* due to milder winters and climate change, increased recreational activity within UK 'hotspots' as shown in figure 1.3 and some links to migration from eastern European countries where there is high prevalence of Lyme borreliosis, in these cases the infection may have not been acquired within the UK (HPA, 2013). Other proposed factors taken from the UK governments guidance for Lyme borreliosis include, increased awareness of the disease, greater access to diagnostic facilities and more sensitive diagnostic methods (HPA, 2013). Many cases of UK Lyme borreliosis are never laboratory confirmed, with the HPA estimating an additional 1000-3000 cases per year and Lyme disease charities suggesting that this number could be even higher at around 15000.

Within the USA Lyme disease has been classed as a nationally notifiable disease since 1991 and alike the UK, cases have rose at a steady pace since 1992 where there were 9908 laboratory confirmed cases which peaked at 30,000 by 2009 (Kuehn, 2013). Under reporting still remains

problematic (Naleway et al., 2002, Coyle et al., 1996) and the CDC estimates that actual cases are likely closer to 300,000 a year (Kuehn., 2013). Using the ten-fold prediction of confirmed to actual cases the CDC has proposed the UK may indeed be encountering more cases than the HPA are predicting and likely in line with estimates made by UK Lyme disease charities. In contrast studies suggest some over-diagnosis may be taking place; Lyme borreliosis symptoms are vague, particularly those associated with post-Lyme disease.

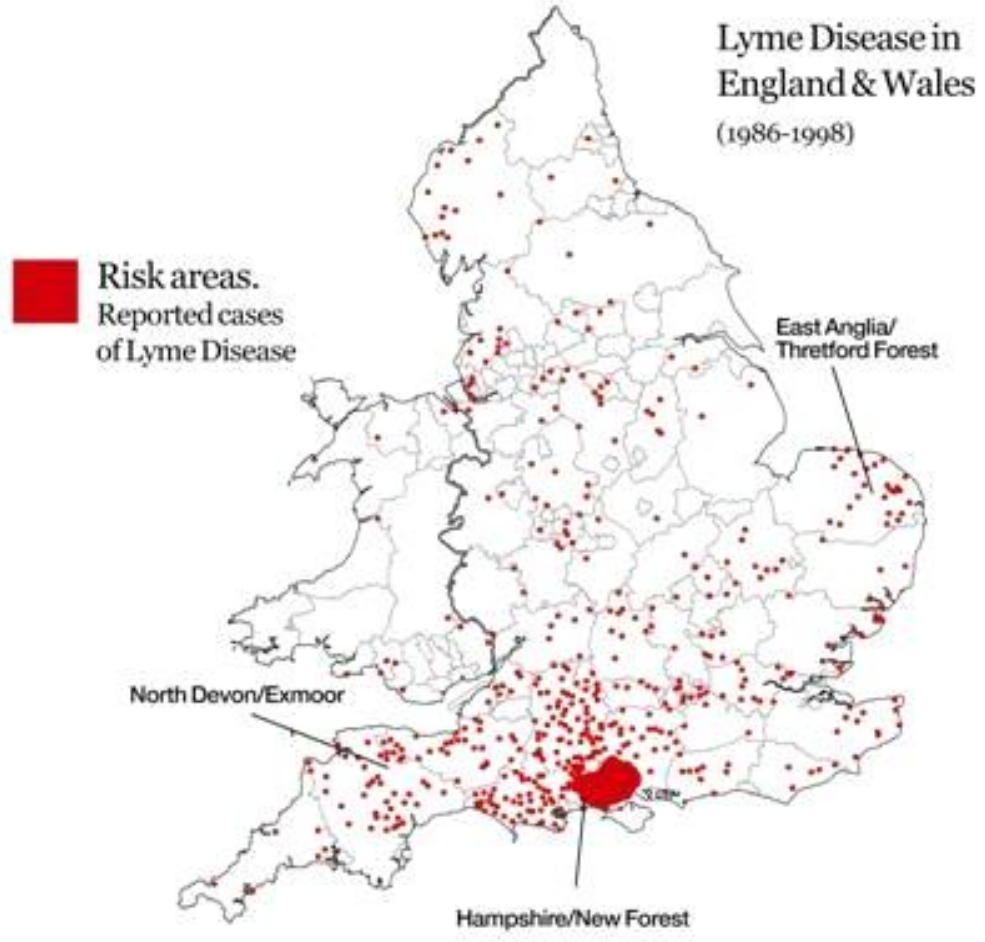


Figure 1.3 - Distribution of Lyme disease in England. A distribution map of Lyme disease cases within England and Wales between 1986 and 1998. During this time the New Forest presented as a 'hot spot' current data suggests these spots have increased to include Thetford forest, Exmoor and the Scottish highlands (Adapted from Smith et al., 2000).

1.1.7. Risk of infection

Age distribution of infection across the globe appears bimodal with spikes between the ages of 5 to 9 years and in the over 50's (Hubálek. 2009, Zeman et al 2013 & Huppertz et al., 1999). Within the UK such rates are similar with the 45-64 age group most at risk of infection (HPA, 2013). All risk groups globally have a fairly evenly spread male to female ratio. It appears that males constitute more in American cases especially within paediatric cases at around 61% (Borchers et al.,2015). However, this is counteracted by the slight predisposition towards females in Europe (Hubálek. 2009 & Zeman et al 2013).

American statistics show that African-American cases are significantly underrepresented (Fix et al., 2000) previous studies have suggested that this may be simply due to demographics with fewer African-Americans living within the endemic rural areas, however a study using data from Maryland observed that fewer infected African-Americans presented with an EM and in some cases this resulted in arthritic late manifestations of the disease (Fix et al., 2000).

The obvious risk factor for *Borrelia* infection is spending recreational time outdoors. However there must be a balance between healthy outdoor activity and the prevention of infection. The most effective way to prevent infection is to avoid tick habitats in endemic areas, such as tall grass and woodland. This however, would rule out some of our much loved British parks and reserves. More sensible precautions include wearing suitable clothing, such as long trousers and socks when walking through long grass, using DEET repellents, the inspection of the body for ticks following exposure and prompt removal of ticks once found. A second risk factor is previous diagnosis of Lyme disease (Bennet et al., 2002, Salazar et al., 1993., Nowakowski et al., 2003), this is likely due to this groups environmental and behavioural risk factor that led to the primary infection being repeated. This risk of secondary infection with Lyme disease also suggests that previous infection may not protect from reinfection, possibly due to variations in circulating strains. Studies on mice demonstrated that immunity could be induced. However, it was strain specific (Barthold, 1996) and current data for Human's propose that this immunity may last $\geq$6 years (Khatchikian, 2014).

Following a tick bite from within an area where Lyme borreliosis is endemic only 3-12% of individuals in Europe (Maiwald et al., 1998, Huegli et al., 2011, Fryland et al., 2010, Nahimana et al., 2004 and Korenberg et al., 1996) and 1-3% in the USA become infected (Sood et al., 1997, Shapiro et al., 1992 and Costello et al., 1989), however upto half of European subjects

seroconvert with no clinical symptoms or classed as asymptomatic where as in the USA asymptomatic seroconversion is fairly rare (Nadelman et al., 2001 & Steere et al., 2003). Interestingly a study of a group of highly exposed Swiss orienteers remained asymptomatic following seroconversion (Fahrer et al., 1991). A recent review article highlighted how remarkable the study was and made the point that over 30% of European patients that present with EM never develop seropositivity (Borchers et al., 2014). Seroconversion in both Europe and the USA remains unclear and much further study is needed.

1.2 Ticks

1.2.1 Ecology, reservoirs and the tick lifecycle

There appears to be distinctive differences between tick vectors across the northern hemisphere and is further discussed in figure 1.4. The hard tick *I. ricinus* is considered the main vector of Lyme disease across Europe with *Ixodes persulcatus* as the most commonly identified vector in Asia (Gray, 1998). The USA demonstrates a split population with *I. scapularis* most commonly found as the spirochetes vehicle in upper midwestern and northeasten areas and *Ixodes pacificus* as the responsible vector in western USA (Gray, 1998).

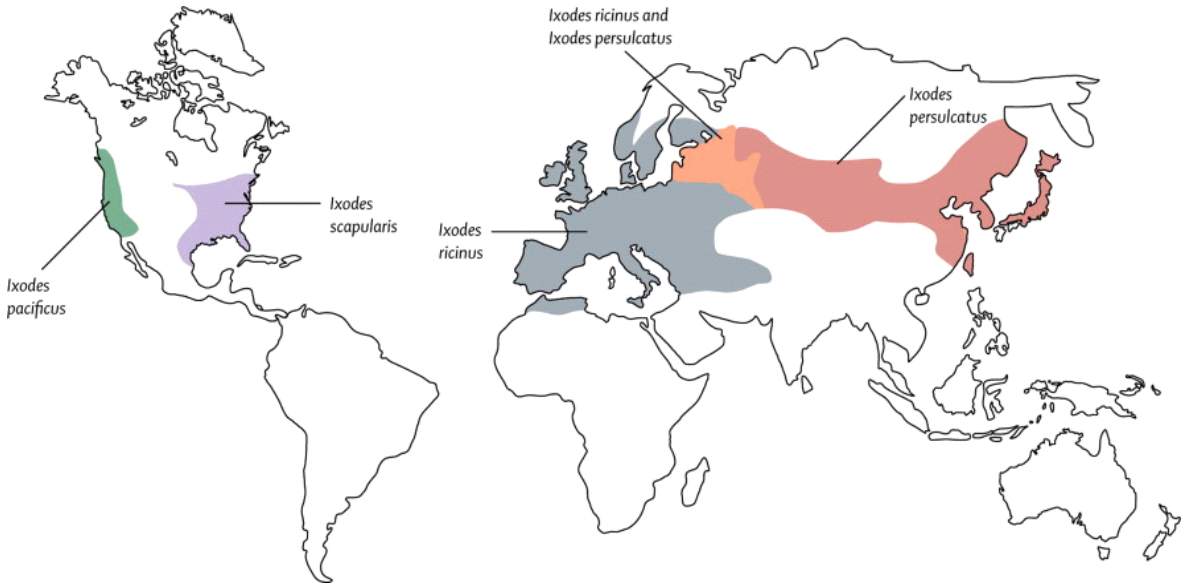


Figure 1.4 - Distribution of hard tick vectors of Lyme *Borrelia* across the world. Within the USA two hard tick vectors *Ixodes scapularis* and *Ixodes pacificus* dominate northeastern and western areas respectively. *Ixodes ricinus* is most prominent in Europe and *Ixodes persulcatus* in Asia. The Baltic states between these areas have a circulation of two *Borrelia* vectors, with a mix of *I. ricinus* and *I. persulcatus* which are both implicated in Lyme Borreliosis. Figure from Stanek et al., 2011.

Females *Ixodes* ticks are predominantly responsible for disease transmission although adult males may ingest host fluids following attachment they rarely withdraw a significant amount of blood (Piesman & Gern, 2005). As a consequence their involvement in disease transmission is considered insignificant although further research is required to confirm this (Piesman & Gern, 2005). The females of all four tick species share a similar four-stage lifecycle where feeding occurs only once during each active stage – egg, larva, nymph and adult, shown in figure 1.5. Post-feeding, the ticks release from the host and return to the soil. If the

environmental conditions are favourable, normally a minimum of 80% humidity, the ticks begin a several month development period before moving into the next stage of the lifecycle (Piesman and Gern, 2005). For adult females this stage results in the laying of around 2000 eggs (Piesman and Gern, 2005). The full length of the tick lifecycle often varies between 2 and 6 years and is largely determined by host availability and the current climate (Piesman and Gern, 2005). Although the *Ixodes* genus is considered as the primary vector for pathogenic *Borrelia* transmission, other ticks have been demonstrated to harbour and transmit the bacteria some of which include *Amblyomma americanum* mainly located in Georgia and Florida (Clark et al., 2013) and *Dermacentor reticulatus* in Spain (Lledó et al., 2014). Much further study is needed to elucidate other competent vectors to gain a better picture of the tick bacteria symbiosis.

Within Europe three species of the *Ixodes* family have been proposed to behave as vectors for *B. burgdorferi s.l.* these include, the previously discussed *Ixodes ricinus* alongside *Ixodes hexagonus* and *Ixodes urinae* (Piesman & Gern, 2005). Although these three species share the same three-stage lifecycle, their ecologies somewhat differ. This vector biology is particularly important and can explain much of a diseases epidemiology. For example, *Ixodes ricinus* is a non-nidicolous tick which occupies open spaces, in contrast *Ixodes hexagonous* and *Ixodes urinae* are nidicolous and are often found in caves and burrows often close to their hosts highlighting a close tick to host relationship (Zeman & Benes, 2013). As non-niducolous ticks occupy an open environment they must actively seek a host and this is often achieved by waiting on the blade edges of long grass. The questing nature of non-nidicolous ticks is responsible for the increased contact they have with humans in comparison to their nidicolous counterparts *Ixodes hexagonous* and *Ixodes urinae* (Zeman & Benes, 2013).

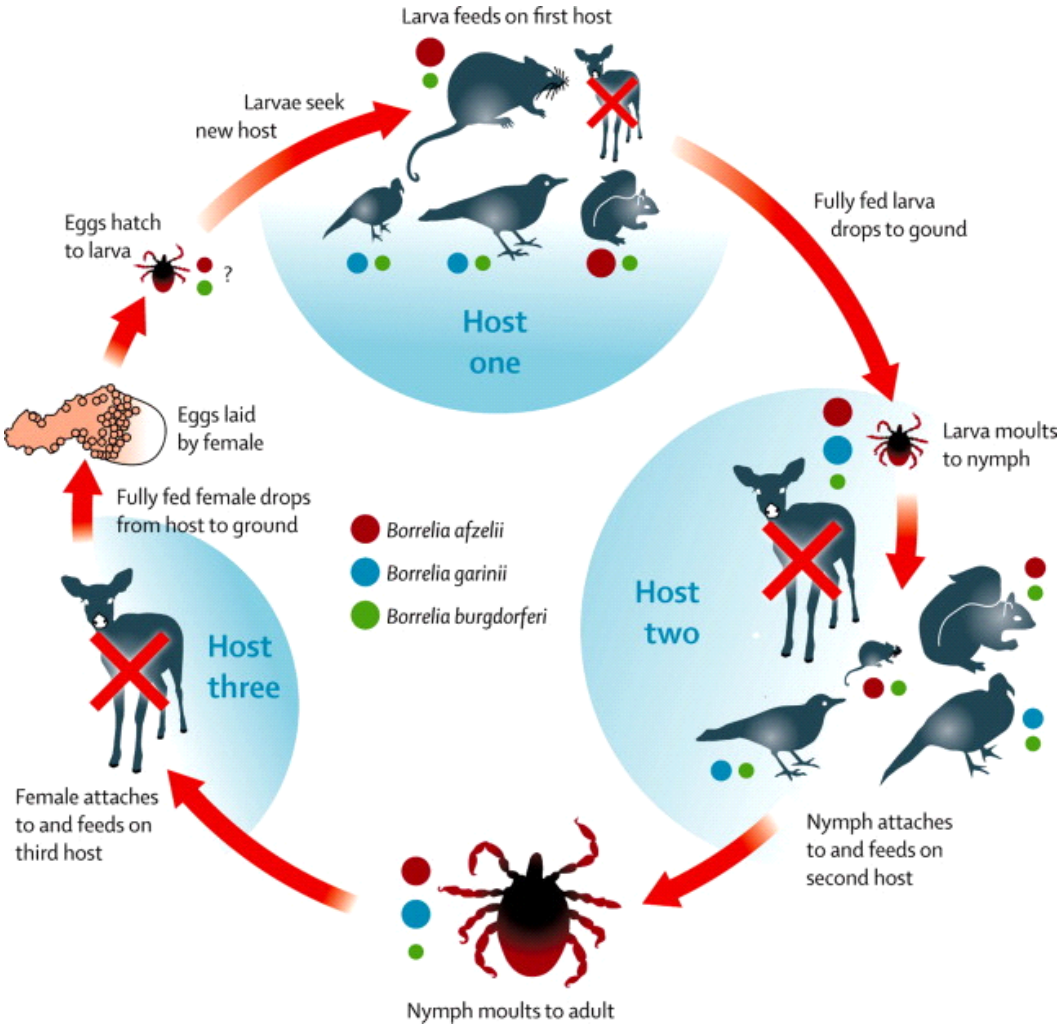


Figure 1.5 - Tick-*Borellia burgdorferi sl* lifecycle. The red crosses indicate non-reservoir hosts. The red dots represent *B. afzelii*, the blue *B. garinii* and the green indicates *B. burgdorferi*. The size of the dots indicate the scale of involvement of each genospecies in each particular organism. Figure from Stanek et al., 2011.

Ticks favour areas with good vegetation cover and a humidity between 80-85% or more simply a mild damp climate (Piesman & Gern, 2005). Current UK hotspots for infection with *Borrelia* include, the New Forest, Thetford, Exmoor, Salisbury Plain, Scottish highlands and islands, the lake district, parts of Sussex, Surrey, Wiltshire, Berkshire and the Yorkshire moors (HPA, 2013). A study of the distribution of *B. burgdorferi s.l.* in *I. ricinus* across central England demonstrated that habitat type significantly influenced the probability of a ticks likelihood of infection (Bettridge et al., 2013). Their findings suggested that ticks found in deciduous woodland were twenty times more likely to be infected with *B. burgdorferi s.l.* than ticks found within moorland. The same study also identified small pockets of central England where *B.*

burgdorferi s.l. infection of *I. ricinus* appeared absent or extremely low, these included North Wales, Lancashire, Eastern Cumbria and Northumberland (Bettridge et al., 2013). However, the study did highlight that although *B. burgdorferi* infection levels were low or absent within these areas it could not be ruled out that these spots may only be temporarily void of infected ticks due to annual fluctuations. In contrast this lack of infection within these areas could be attributed to habitat type which in effect could influence tick questing.

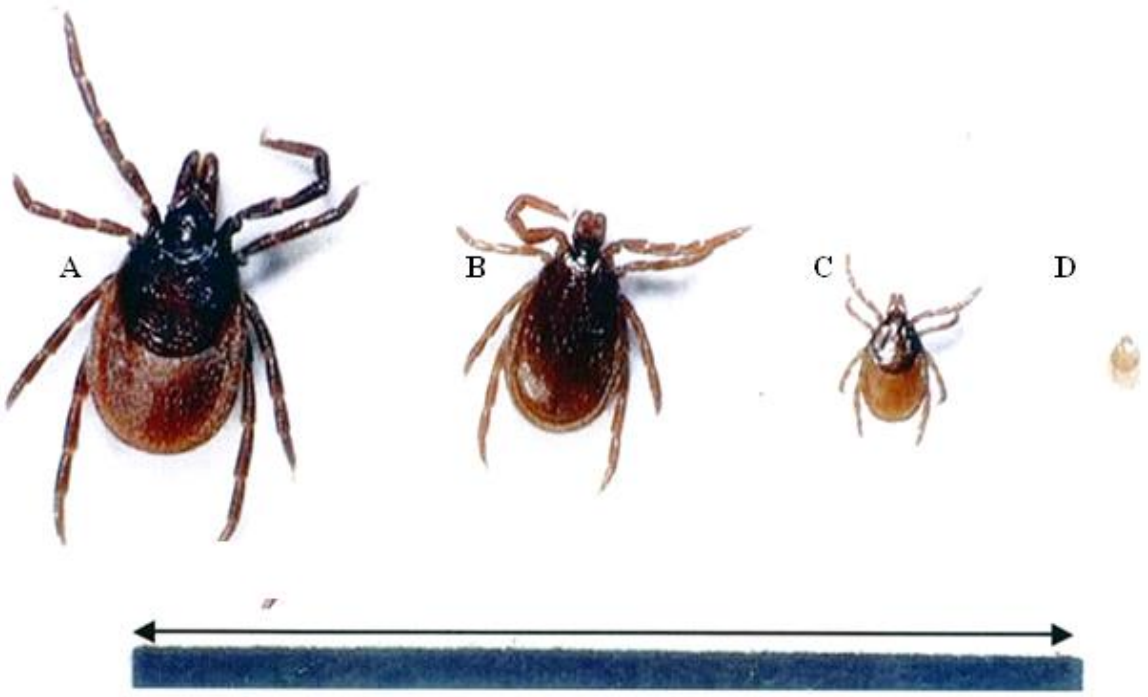


Figure 1.6 - *Ixodes ricinus* ticks at various stages of the lifecycle. A. B. C. D. The blue bar represents 1.5cm. Figure adapted from a photograph taken from Phillippe Parola Collection (Extract from the 16th Consensus Conference on Anti-infective therapeutics 2006).

1.2.2 Tick morphology

Ixodes ticks have a relatively large body of around 2-30mm and have 4 pairs of legs as a nymph and adult. The gnathosoma or the mouthparts comprise of the hypostome, palps and chelicerae which are clearly visible in all stages with backwardly pointing teeth on the hypostome shown in figure 1.7. Ticks are believed to quest for a host at the tip of a blade of grass by waving their front legs in the direction of an approaching host in order to aboard should the animal brush past (Cook, 2015). The setae within the Haller's organ and the tip of the front legs can detect host stimuli such as vibrations, carbon dioxide, humidity and temperature allowing preparation to claim a new host.

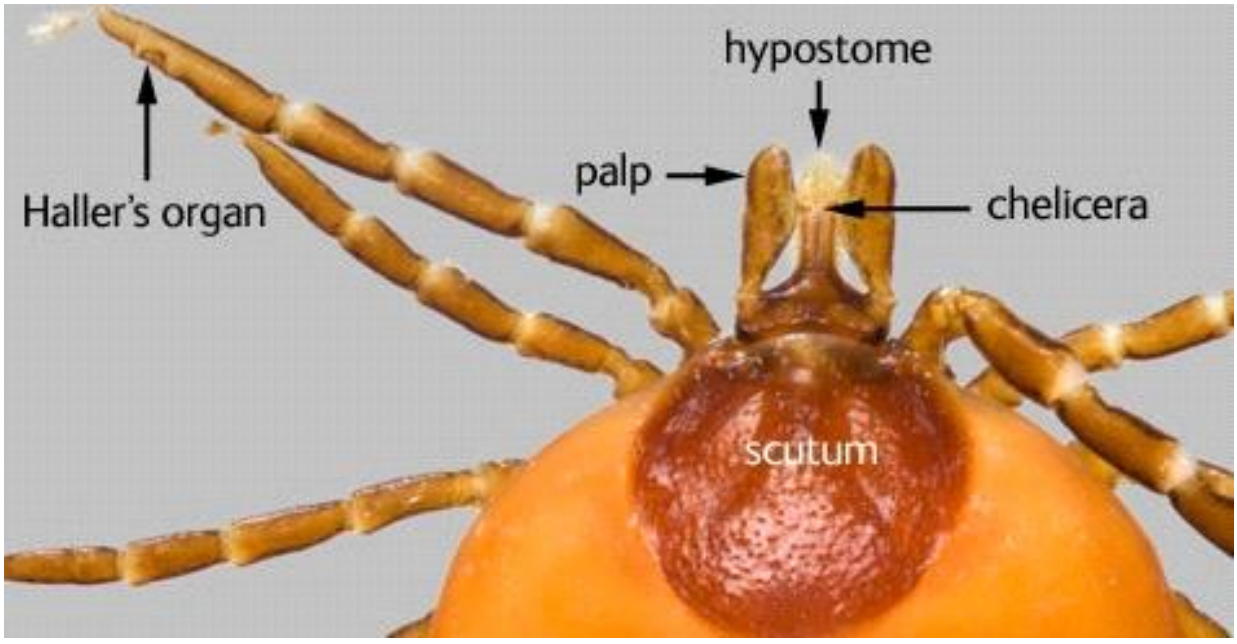


Figure 1.7 - The anatomy of an Ixodes ricinus tick. Credit to Peter York – National History Museum London. The *Ixodes* mouthparts or gnathosoma are composed of the hypostome, palps and chelicerae. The Haller's organ is located at the tip of the front legs and is essential for the detection of host stimuli.

1.2.3 Feeding of ticks and the transmission of *Borrelia*

As a general rule feeding occurs for around 3 days for larvae, 5 for nymphs and 7 days for adult females. (Stanek et al., 2011). After locating a host the tick begins to attach using numerous projections on the hypostome, these hook type appenditures anchor the tick following penetration of the skin by the hypostome (Cook, 2015). Adult ticks can become particularly engorged during a blood meal with their weight increasing considerably compared to their native size (Ribeiro et al, 2006) Spirochaetes including *Borrelia* can reside in the midgut of

unfed ticks and begin to replicate and disseminate to the salivary glands during a blood meal. (Ribeiro, 1987).

In order for a tick to successfully obtain a blood meal and for *Borrelia* to infect a new animal a number of host defences must be confronted. A humans first line of defence is the skin which the tick penetrates via the hypostome giving access to the host blood supply. Tick saliva provides the first contact and contains a mixture of factors which alleviate host pain, ensure a continual blood supply to the area of penetration and prevent coagulation (Tyson et al., 2007, Chmelar et al., 2012 and Ribeiro, 1987).

Animal models suggest that infection with *Borrelia* rarely occurs within 24 hours of tick attachment with effective transmission peaking between 48-72 hours (Dattwyler et al., 2005, Kalish et al., 2001, Steere and Angelis, 2006, Lakos and Solymosi, 2009). Some experimental data have demonstrated that Bb-infected *I. scapularis* nymphs can transfer a significant number of spirochaetes as early as 24 hours following attachment although further studies are required (Weber et al., 1990). It is likely that the time period required for effective transmission is both strain and host species specific (Dattwyler et al., 2005, Kalish et al., 2001). Transmission speed of *I. ricinus* within Europe doesn't appear to follow the same trend. Data suggest that the host can become infected with *B. afzelii* as early as 17 hours following attachment however the current evidence is limited and based upon the gerbil model (Klempner et al., 2001, Krupp et al., 2003). The study also noted that *B. afzelii* was more rapidly transmitted by *I. ricinus* than the transfer of *B. burgdorferi* which highlights the difficulty in giving accurate guidance about the relationship between attachment time and risk.

1.3 *Borrelia*

1.3.1 Spirochaetes and *Borrelia*

Spirochaetes are small, coiled bacteria which are often responsible for complex diseases. The family covers *Borrelia, Leptospira, Serpulina* and *Treponema*. Certain strains of *Leptospiridia* are the causative agent of Weil's disease a rare systemic infection which can lead to renal failure and liver damage (Inoue et al., 2010). *Treponema pallidum* is the bacteria responsible for syphilis and other family members have been associated with periodontitis and endocarditis (Golden et al., 2003). Of the currently known *Borrelia* strains almost one third can cause Lyme disease (Hengge et al., 2003), a further proportion are responsible for relapsing fever and the remaining strains are uncharacterised in terms of their ability to colonise a human host.

1.3.2 The genus *Borrelia*

The genus *Borrelia* was principally identified in 1834 and so named *Spirillium giganteum* which was later changed to *Borrelia* in 1907 (Wright, 2009). There are over thirty species of *Borrelia* described within scientific literature, with some newer species identified over recent years (Margos et al., 2011 and Rudenko et al., 2010*). B. burgdorferi sensu stricto, B. afzelii* and *B. garinii* are collectively known as *B. burgdorferi sensu lato* and are the main species responsible for Lyme borreliosis within humans, (Rizzoli et al., 2011) while *B. duttonii, B. hermsii and B. reurrentis* are most commonly associated with relapsing fever (Schwan et al., 2003). A recent phylogenomic study has proposed that the genus *Borrelia* should be divided into two genera. This recommendation was that all relapsing fever *Borrelia* should belong to the genus *Borrelia* and the creation of a new genus, *Borreliella gen. nov.* which would contain all Lyme disease causing *Borrelia* (Adeolu and Gupta, 2014). The three species that will be discussed throughout the research project include, *B. afzelii, B. garinii and B. burgdorferi s.s.* and from here on denoted as *B. burgdorferi* s.l. Any reference to the term *Borrelia* accounts as a generalisation to both relapsing fever and Lyme disease causing spirochaetes.

1.3.3 Relapsing Fever *Borreliae*

As previously discussed spirochaetes from the genus *Borrelia* are also responsible for the bacterial infection termed relapsing fever. This infection most commonly manifests as recurring episodes of flu-like symptoms such as fever, headaches, muscle and joint pain and nausea (Dworkin et al., 2008). Relapsing fever is often separated into two branches consisting of tick-borne relapsing fever (TBRF) and louse-borne relapsing fever (LBRF).

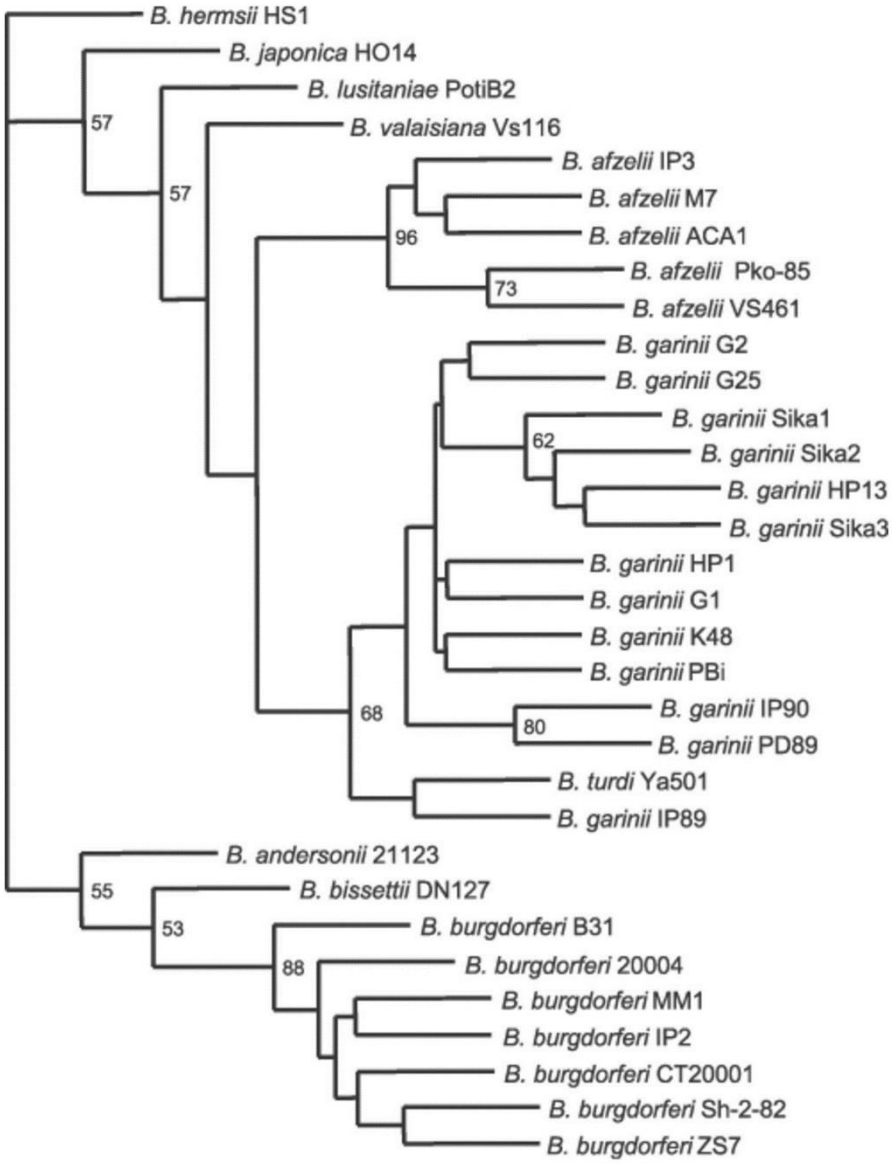


Figure 1.8 - Phylogenetic tree generated from *groEL* gene sequences from *B. burgdorferi sl* strains. The phylogenetic tree was taken from Lee et al., 2003. The tree was constructed using the maximum-likelihood method. The *groEL* gene was selected as amino acid sequences within this frame are highly conserved across *Borrelia* species.

Alike *Borrelia* strains associated with Lyme disease, the number of *Borrelia* species associated with relapsing fever has almost doubled since the 1980s. Over 17 strains are now associated with relapsing fever with a number of others awaiting further characterisation and isolation (Cutler, 2006). Within ticks, relapsing fever is mainly transmitted by the *Ornithodoros* genus using a similar mechanism as what is seen with the transmission of Lyme disease by *Ixodes* ticks. However relapsing fever *Borreliae* have also been demonstrated to be transmitted by louse, potentially harnessing the spirochaete with epidemic potential (Cutler, 2006).

Genospecies	Distribution	Human disease manifestations
B. afzelii	Europe	Lyme disease
B. americana	North America	Lyme disease
B. andersonii	North America	Lyme disease
B. bavariensis	Europe, Asia	Lyme disease
B. bissettii	Europe, North America	Lyme disease
B. burgdorferi	Europe, North America	Lyme disease
B. californiensis	North America	Lyme disease
B. carolinensis	North America	Lyme disease
B. crocidurae	Africa	Relapsing fever
B. duttonii	Africa	Relapsing fever
B. finlandensis	Europe	Lyme disease
B. garinii	Europe, Asia	Lyme disease
B. hermsi	North America, Canada	Relapsing fever
B. hispanica	Africa	Relapsing fever
B. japonica	Japan	Lyme disease
B. kurtenbachii	North America	Proposed to be involved in Lyme disease
B. lonestari	North America	Southern tick-associated rash illness (STARI)
B. lusitaniae	Europe	Lyme disease
*B. miyamotoi**	North America, Japan, Europe	Unknown
B. parkeri	North America, Canada	Relapsing Fever
B. recurrentis	Africa	Relapsing Fever
B. sinica	China	Lyme disease
B. spielmanii	Europe	Lyme disease
B. tanuki	Japan	Prevalence in humans unconfirmed
B. turdi	Japan	Prevalence in humans unconfirmed
B. turicatae	North America, Canada	Relapsing Fever
B. valaisiana	Europe, Asia	Lyme disease
B. yangtze	China	Prevalence in humans unconfirmed

Table 1.3. A summary of known *Borrelia* strains and their global distribution. Table compiled using literature data, searching for strains isolated from human hosts. **B. miyamotoi* – The human disease manifests itself as a 'Lyme disease-like' syndrome without the characteristic EM rash. The strain is genetically related to relapsing fever spirochaetes with the first human infection reported in Russia in 2011, the spirochaete has now been identified across both North America, Asia and Europe.

1.3.4 Genomic complexity

The genome of *B. burgdorferi* can be considered as one of the most complex arrangements seen within all bacteria (Casjens et al., 2000, Casjens et al., 2012, Fraser et al., 1997) and is made up of a ~950kb linear chromosome and a variety of both circular and linear plasmids which range in size from 9 to 62kb (Brisson et al., 2012). In total *Borrelia* are known to carry the largest number of plasmids ever described in bacteria (Reyes-Lamothe et al., 2012). The linear chromosome is highly conserved across many *Borrelia* species (Casjens et al., 1995, Fraser et al., 1997, Glöckner et al., 2004, Glöckner et al., 2006, Schutzer et al., 2010). However, extensive variation is observed among the plasmid elements, an example is the highly variable OspC protein which is encoded on the 26kb plasmid. In general all of the plasmids are maintained at a low copy number, (Beaurepaire and Chaconas, 2005), with the exception of a small circular plasmid isolated from the Ip90 strain which seems to have a higher copy number (Dunn et al., 1994). Within the *Borrelia* family such plasmids are often referred to as essential genetic elements, (Stewart et al., 2005) or mini chromosomes (Barbour and Zückert, 1997) as some contain genes which are significant regulators of the tick-vertebrate lifecycle (Norris et al., 2010). The linear counterparts have covalently closed telomeres (Barbour and Garon, 1987) and replication is dependent upon ResT, an essential telomere resolvase (Kobryn and Chaconas, 2002).

The majority of housekeeping genes are found upon the linear chromosome and although some plasmids are essential for propagation of the enzootic cycle they are not required for *in vitro* culturing (Stewart et al., 2005). One particular single copy plasmid, Cp26 is the only known essential plasmid for *in vitro* growth. This plasmid encodes the ResT enzyme essential for the replication of *Borrelia's* linear DNA (Kobryn and Chaconas, 2002). Curiously the same plasmid also encodes the outer surface protein C (OspC) which is crucial for successful transmission from the tick to the vertebrate host (Grimm et al., 2004, Pal et al., 2004, Stewart et al., 2006, Tilly et al., 1997).

Although members of the genus *Borrelia* are considered to possess the most complex genomes of all known bacteria, their genome is often defined as vastly redundant, due to its particularly large number of paralogous genes, specifically among the cp32 family of plasmids, where large stretches of fundamentally identical genes can be located and are interrupted by variable genes encoding lipoproteins (Casjens et al., 2011). The first *B.burgdorferi* strain sequenced was B31-MI and was found to be composed of a single linear chromosome alongside 12 linear and 9

circular plasmids, which contribute approximately 47% to the total genome as shown in figure 1.9. This unusual makeup of *Borrelia* makes genetic manipulation difficult, plasmid loss during experiments is one of the main challenges during *in vitro* experiments (Brisson et al., 2012). Many of *Borrelia's* proteins remain with unknown function which is further complicated by the fact that around 30% of *Borrelia's* open reading frames share no sequence similarity to previously annotated genes (Casjens et al., 2000).

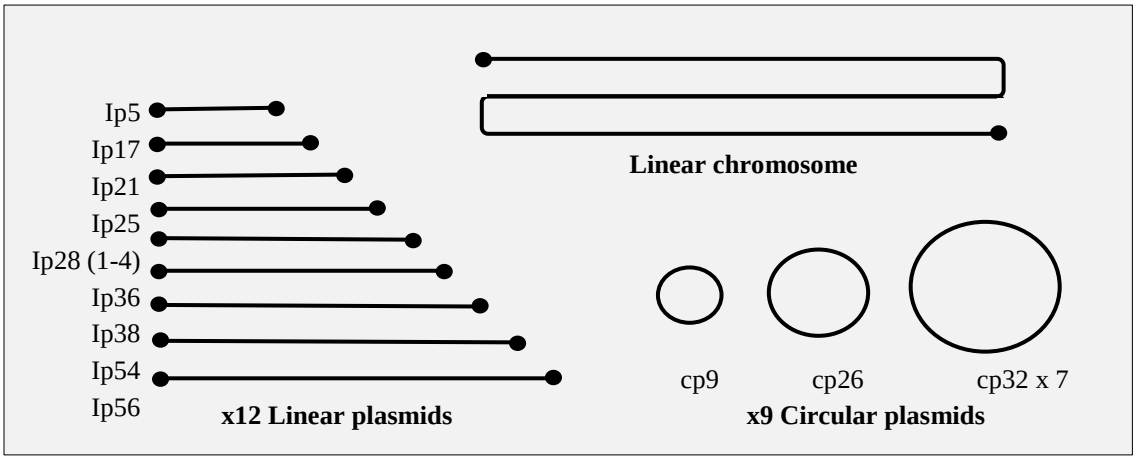


Figure 1.9 Genetic arrangement of *B. burgdorferi* B31 strain. The *B. burgdorferi* strain B31 is comprised of a single linear chromosome, 12 linear plasmids and 9 circular plasmids. All of the linear components contain covalently closed telomeres and require the ResT enzyme for replication of the linear DNA. Figure created using genomic data from Casjens et al., 2000.

A further oddity *of B.burgdorferi* is its absence of many metabolic pathways. The bacteria is an auxotroph for all nucleotides, amino acids and fatty acids, it also lacks genes encoding essential enzymes for oxidative phosphorylation and the tricarboxylic acid cycle and instead acquires energy via the fermentation of sugar to lactic acid (Fraser et al., 1997). The lack of such pathways and metabolic enzymes is possibly explained by reductive evolution, as the organism moved towards a lifecycle parasitizing required nutrients from its hosts. Curiously much research suggests *B.burgdorferi* does not require iron (Posey & Gherardini, 2000 & Troxell et al., 2012), normally essential for many metabolic activities of bacteria (Messenger & Barclay, 1983). To date no genes have been identified within the spirochaetes genomes that are known to encode iron-requiring proteins such as *Borrelia's* single superoxide dismutase (SOD), which is now thought to utilise manganese as its cofactor (Troxell et al., 2012).

1.3.5 Morphology and Ultrastructure

In terms of morphology, *Borrelia burgdorferi s.l.* can be described as coiled, highly motile spirochaetes with a length of 20-30µm (Krupka et al., 2008). The species was originally classified as Gram-negative due to the identification of both an outer and cytoplasmic membrane. However, the group showed distinctive differences from their counterparts resulting in the creation of a new phylum (Paster et al., 1991). *Borrelia* harness a unique mode of motility by means of their endoflagella or axial filaments, which enable successful survival in viscous environments such as intestinal tracts, mucosal tissues and possibly biofilm (Guttenplan & Kearns, 2013). In contrast to bacteria with an extracellular flagella, a spirochaetes endolflagella is attached to both termini of the cytoplasmic membrane as shown in figure 1.10. This endoflagella spans the entire cell essentially weaving from one side of the protoplasmic cylinder to the other. Within *B. burgdorferi s.l.*the number of these can be anywhere between 7 and 11 (Hovind-Hougen, 1984 and Motaleb et al., 2000) and are the principle component responsible for the bacteria's unusual shape (Motaleb et al., 2000).

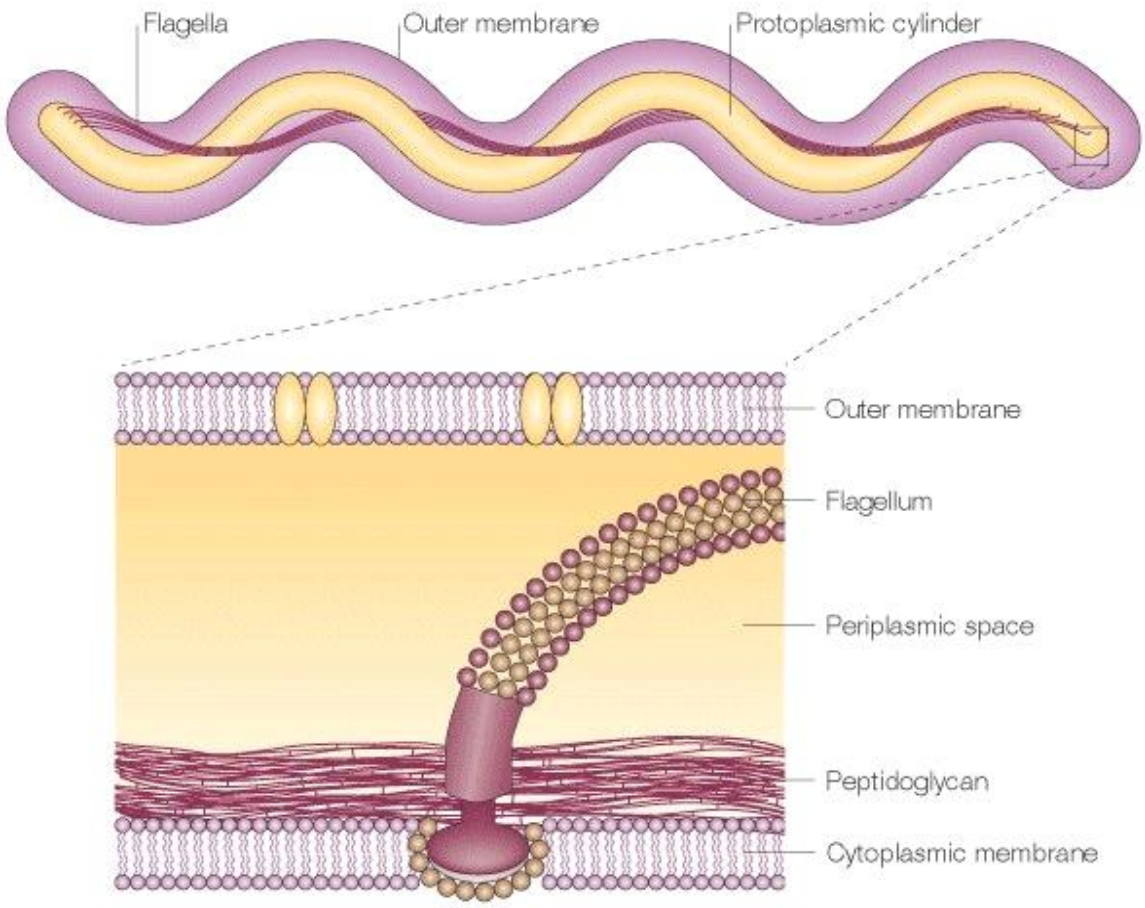


Figure 1.10. The ultrastructure of the spirochaete *Borrelia*. Figure taken from Rosa et al., 2005. The bacterium is enveloped by both an outer and a cytoplasmic membrane. Between these two structures and within the periplasmic space exists a layer of peptidoglycan. Multiple flagella span the periplasmic space and are attached at the cytoplasmic membrane at each termini. These flagella weave around the protoplasmic cylinder and are responsible for the spirochaete's unusual shape (Rosa et al., 2005).

1.3.6 Alternative morphological states

It has been known for a number of years that *B. burgdorferi s.l* is capable of initialising drastic changes in its morphology (Barbour & Hayes, 1986, Miklossy et al., 2008, Garon et al., 1991) however the reasons why this occurs and whether this is of clinical importance remains unclear. Many *Borrelia* strains including *B. burgdorferi s.l.* have the capacity to produce extracellular vesicles or blebs when grown *in vitro* (Barbour & Hayes, 1986, Garon et al., 1991) and there is some evidence that these can occur during infection (Dorward et al., 1991). Cryoelectron tomography studies have revealed that these outer membrane vesicles are often released close to areas of cell division (Kudryashev et al., 2009) and exposure to stress appears to be enough

to initiate their production and release from the membrane. These blebs are known to contain, plasmid DNA, OspA and OspB proteins alongside a significant number of other cytosolic and inner membrane molecules including the *Borrelia* protein enolase which is responsible for plasminogen binding and activation (Toledo et al., 2011). Curiously the vesicles must also carry some of these adhesion proteins as exposed components upon the outer membrane as the vesicles have been demonstrated to bind to human endothelial cells (Shobery & Thomas, 1993).

Cyst formation has also been described for *Borrelia* and involves a huge shift in the structure of the cells shape and changes to membrane integrity (Alban et al., 2000). Research suggests (Alban et al., 2000) that cyst formation occurs in response to starvation and can be induced *in vitro* by the cultivation of *Borrelia* in serum free BSKII media (Alban et al., 2000). Many atypical forms of *Borrelia* have been described and include cystic, loops, rings and end knob forms (Miklossy et al., 2008). Although the process is particularly slow it is thought to still be an active process and requires some level of continuous protein synthesis (Bergström & Zückert, 2010). The development into a cystic form also appears to be a reversible action with studies suggesting upto 53% of cells can then revert back to the active coiled morphology (Bergström & Zückert, 2010). The implications of cyst formation *in vivo* still remains unclear and needs much further research in order to understand the complex cycle this unique bacterium participates in.

1.4 The *Borrelia* outer membrane

1.4.1 Introduction to the outer membrane and its diagnostic and therapeutic importance

The components of the outer membrane of all bacteria play a major role in the survival of the micro-organism. These components are particularly important when the external environment of the micro-organism is frequently changing, for example when the bacteria switch host. This outer membrane must function appropriately and maintain control of influx and efflux whilst remaining stable. Proteins and cell membrane components play a crucial role in ensuring the system is not breached allowing the bacteria to adapt to the ever changing environment. Within *E. coli* approximately 50% of the mass of a bacterium's outer membrane consists of protein whether this be in the form of lipoproteins or integral membrane proteins (Koebnik et al., 2000) highlighting the importance of these exterior membrane features.

The membranes found with *B. burgdorferi* are often described as unique and consist of a variety of glycolipids, a huge array of surface exposed lipoproteins and a smaller number of integral membrane proteins (Kudryashev et al., 2009), yet lack highly inflammatory lipopolysaccharides (LPS) often found within Gram-negative bacteria (Takayama et al., 1987). Two of these glycolipid components contain cholesterol and include ACGal (cholesteryl 6-O-acyl-β-D-galactopyranoside) and CGal (cholesteryl-β-D-galactopyranoside), whilst the third. MGalD (mono-α-galactosyl-diacylglycerol) does not (Ben-Menachem et al., 2003 and Schroder et al., 2003). Curiously the *B burgdorferi* genome is not known to encode any pathways for cholesterol biosynthesis which is further supported by the fact that *B. burgdorferi* cannot be grown *in vitro* unless the media is supplemented with cholesterol (Coleman et al., 2015). As *B. burgdorferi* cannot produce its own supply of these biomolecules and yet cannot sustain life without them it has been presumed that the spirochaete must acquire cholesterol from the tissue and fluids from the infected host. This assumption has been further supported by recent evidence for a novel mechanism *in vitro* where by *B. burgdorferi* can acquire cholesterol by contact with the plasma membrane of eukaryotic cells (Crowley et al., 2013). A further oddity of the spirochaetes relationship with cholesterol is that cholesterol glycolipids can amalgamate to form lipid rafts within the outer membrane of *B. burgdorferi* which although are often seen within eukaryotic cells they are poorly understood and recognised within prokaryotes. Within eukoryotic membranes these lipid rafts have been associated with vesicle formation (Huttner and Zimmerberg, 2001), cytosis (Chen and Rand, 1997), membrane elasticity (Salaün et al., 2004) and the lateral sorting of proteins and receptor clustering (Epand 2008).

Borrelia's outer surface provides the platform between itself and the infected host, highlighting that proteins associated on the outer membrane must hold some responsibility for *Borrelia's* capacity to disseminate throughout the host tissues. Alongside their roles in dissemination these surface molecules must also provide mechanisms for the spirochaete to remain hidden from the immune response. The arrangement of such proteins within the outer membrane appears to vary in line with the spirochaetes enzootic cycle, its host and importantly in accordance to its environment (Fraser et al., 1997, Babb et al., 2001). There are several protein families which have been identified to be associated with *Borrelia's* outer membrane and include a significant number of lipoproteins (~100) such as, outer surface proteins (osp's), decorin binding proteins and complement regulator acquiring proteins (CRASPs) (Skare et al., 1997, Cassatt et al., 1998) and a smaller number of integral membrane spanning proteins. The wide array of lipoproteins is likely to be one of *Borrelia's* tools providing on-going immune evasion strategies allowing successful dissemination throughout the host. A full list of key *Borrelia* surface proteins are summarised in table 1.4.

Over the past decade several reviews have highlighted the importance of identifying and characterising bacterial outer surface proteins with the hope of identifying new drug and vaccine targets. This has been particularly noted for *Borrelia* where many proteins remain annotated as hypothetical and hold no known function (Kenedy et al., 2012). As *B. burgdorferi* is considered an extracellular pathogen, much research has focused on the production of a vaccine allowing humoral immunity to protect any given individual from active disease. During 1998 the outer surface protein A (OspA) based vaccine was licensed and approved for human use. Although this vaccine was commercially available for several years it is no longer in use. The withdrawal of the OspA based vaccine has directed much research towards the identification of *Borrelia* antigens that fulfil the three requirements of a good vaccine candidate, these include,

i. The antigen is surface exposed.
ii. The antigen is conserved among several strains.
iii. The antigen is produced both during tick transmission and mammalian infection.

Alongside the importance of identifying new antigens, proteins situated on the outer membrane are at the direct interface between the host and the invading pathogen, they often have crucial cellular roles with some moonlighting as important virulence factors with examples summarised in table 1.4.

Protein	Alternative names	Type	Known functions	Expression	Reference
OspA	BBA15	Lipoprotein	Adhesion to tick gut Spirochaete-spirochaete aggregation Degradation of the ECM Adhesion to host cells Activation of host immune cells	Vector/Host	Pal et al., 2004, Pat et al., 2000, Fuchs et al., 1994, Pulzova et al., 2011.
OspB		Lipoprotein	Adhesion to tick gut	Vector	Fikrig et al., 2004
OspC		Lipoprotein	Adhesion to tick salivary glands Establishment of early host infection Degradation of the ECM	Vector/Host	Ramamoorthi et al., 2005, Pal et al., 2004, Lagal et al., 2006.
OspD		Lipoprotein		Vector	Li et al., 2007.
OspF		Lipoprotein			Lam et al., 1994.
VlsE		Lipoprotein			Eicken et al., 2002
P66	Oms66, BBO603	OM spanning	Channel forming porin	Host	Skare et al., 1997
BBK32		Lipoprotein	ECM adhesion	Vector/Host	Fikrig et al., 2000
CRASP-1	BBA68	Lipoprotein	Complement system evasion ECM degradation Adhesion to host cells Adhesion to the ECM	Host	Bykowski et al., 2008, Haupt et al., 2007, Hallstrom et al., 2010, Gilmore et al., 2008.
CRASP-2	BBH06	Lipoprotein	Complement system evasion	Host	Kraiczy et al., 2001, Brissette et al., 2009.
CRASP-3	ErpP	Lipoprotein	Complement system evasion ECM degradation	Host	Schwab et al., 2013, Haupt et al., 2007, Siegel et al., 2010, Brissette et al., 2009.
CRASP-4	ErpC	Lipoprotein	Complement system evasion ECM degradation	Host	Schwab et al., 2013, Haupt et al., 2007, Brissette et al., 2009.
CRASP-5	ErpA	Lipoprotein	Complement system evasion ECM degradation	Host	Haupt et al., 2007, Siegel et al., 2010, Brissette et al., 2009, Schwab et al., 2013.
DbpA	BBA24	Lipoprotein	Adhesion to ECM via decorin		Fischer et al., 2003
DbpB	BBA25	Lipoprotein	Adhesion to ECM via decorin		Fischer et al., 2003
RevA			Adhesion to the ECM via Fibronectin and laminin	Host	Brissette et al., 2009
RevB			Adhesion to the ECM via fibronectin	Host	Brissette et al., 2009
BmpA			Purine transport	Host	Verma et al., 2009

			Adhesion to the ECM via laminin		
BmpB			Adhesion to the ECM via laminin	Host	Verma et al., 2009
BmpC			Adhesion to the ECM via laminin	Host	Verma et al., 2009
BmpD			Adhesion to the ECM via laminin	Host	Verma et al., 2009
ErpX		Lipoprotein	Adhesion to the ECM via laminin	Host	Brisette et al., 2009.
Bgp	BBO588		Adhesion to the ECM via glycosamino glycan and heparin	Host	Parveen et al., 2006, Cluss et al., 2004.
BBB07			Adhesion to the endothelium via α3β1 and integrins	Host	Behera et al., 2007
BB0690			Persistence during tick starvation	Vector	Li et al., 2007
BptA	BBE16		Persistence within the tick	Vector	Revel et al., 2005
BBE31			Tick gut/salivary glands migration vua TRE31	Vector/Host	Zhang et al., 2011
BBA52			Tick gut/salivary glands migration	Vector/Host	Kumar et al., 2011
BBE22	PncA Nicotinamidase		Tick gut/salivary glands migration	Host	Purser et al., 2003
BBA64	P35		Tick gut/salivary glands migration Persistent host infection	Vector/Host	Gilmore et al., 2007, Gilmore et al., 2008
BBA65			Tick gut/salivary glands migration	Vector/Host	Gilmore et al., 2007
BBA66			Tick gut/salivary glands migration	Vector/Host	Gilmore et al., 2007
BBA73			Persistent host infection	Host	Hughes et al., 2008
BBA03			Tick/host transmission	Vector/Host	Bestor et al., 2012
BBO250			Cell division and envelope integrity	Vector/Host	Liang et al., 2010
BBA07			Tick/host transmission	Vector/Host	Xu et al., 2010
BBA74	Oms28		Tick environment adaption	Vector/Host	Mulay et al., 2009, Mulay et al., 2006
FliG1			Motility in viscous media	Vector/Host	Li et al., 2010
BesC	BB0142	OMP	Porin		Bunikis et al., 2008
DipA	BB0418	OMP	Dicarboxylate-specific porin		Thein et al., 2012
BamA	BB0603	OMP	Barrel assembly machinery	Vector/Host	Lenhart et al., 2010

Table 1.4. A summary of *B.burgdorferi* surface proteins with known functions and corresponding ligands. Adapted from Pulzova and Bhide 2014.

1.4.2 Lipoproteins and *Borrelia*

As the name suggests 'lipoproteins' may be described as any conjugated protein whereby the non-protein moiety consists of at least one lipid; this lipid may be cholesterol, phospholipid, triacylglycerol or a combination of the above. Bacterial lipoproteins cover a broad range of functions including cellular physiology (Sutcliffe et al., 1995) and replication (Alloing et al., 1994) with many significantly contribute to bacterial virulence (Cassatt et al., 1998, Haake. 2000, Schroder et al. 2003). Across both Gram-positive and Gram-negative bacteria lipoproteins have been implicated in adhesion, conjugation, nutrient uptake, sporulation and the extracytoplasmic folding of proteins (Alloing et al., 1994, Mathiopoulos et al., 1991, Perego et al., 1991, Sutcliffe et al., 1995). Within pathogenic bacteria these roles extend to key virulence factors including, colonisation, evasion of the host defences and immunomodulation (Cassatt et al., 1998, Haake. 2000, Schroder et al. 2003).

Bacterial lipoproteins are initially translated as preprolipoproteins with a characteristic secretion signal peptide. This N-terminal signal peptide, of usually around 20 amino acids is present in the biosynthesis of both Gram-positive and Gram-negative bacterial lipoproteins (Inouye et al., 1977). Lipoprotein diacylglyceryl transferase (Lgt) is then responsible for the modification of a conserved sequence at the C-terminal of the signal peptide known as the lipobox, [LVI][ASTVI][GAS]C. Lgt catalyses the covalent attachment of a diacylglycerol moiety to the thiol group on the side chain of the cysteine residue resulting in the production of prolipoprotein (Babu et al., 2006). Following lipidation the signal sequence of the prolipoprotein is cleaved by lipoprotein signal peptidase which leaves the cysteine of the lipobox as the new amino-terminal residue (Tokunaga et al., 1982).

Within Gram-negative bacteria and GC-rich Gram-positive bacteria the newly cleaved prolipoprotein undergoes an additional modification (Vidal-Ingigliardi et al., 2007). Lipoprotein N-acyl transferase (Lnt) catalyses the attachment of an amide-linked acyl group to the N-terminal cysteine residue. However, lipoproteins produced both and with and without this additional modification anchor tightly to the membrane following signal peptide-dependent translocation across the cytoplasmic membrane (Vidal-Ingigliardi et al., 2007). The structure of the acyl chains are believed to be predominantly derived from membrane phospholidids for *Borrelia* these include phosphatidylcholine and phosphatidylglycerol. However, studies suggest that there may be variations in the structures of the acyl chains with one interesting case from *B.burgdorferi* (Bouchon et al., 1997) where the lipid component of

their lipoproteins contains an acetyl group in replacement for one of the ester linked fatty acids (Beermann et al., 2000). As many lipoproteins are surface exposed it has been questioned whether such acyl moieties play other roles alongside anchorage to the membrane (Kovacs-Simon et al., 2011).

As previously addressed bacterial lipoproteins represent a significant number of virulence factors, they are particularly important within the spirochaete genus and are the most abundant proteins that are expressed by all spirochaetes (Haake. 2000, Liang et al., 2002, Radolf et al., 1994). *B.burgdorferi* is known to encode over 100 lipoproteins (Fraser et al., 1997) with approximately 8% of all genes likely to fall within this category (Casjens et al., 2000).

1.4.3 Bacterial integral membrane proteins

1.4.3.1 Introduction to beta-barrel proteins

Outer membrane proteins (OMPs) make up a smaller part of the outer membrane of gram-negative bacteria in comparison to those proteins which are anchored to the membrane by N-terminal lipids. OMPs are found to span the outer membrane of all Gram-negative bacteria, chloroplasts and mitochondria (Schulz, 2000). To date all OMPs have been identified or characterised as β-barrels. These integral proteins come in various sizes and have been demonstrated to span the outer membrane from 8 to 22 times (Khalid et al., 2006) and follow a set assembly format which is further discussed in Schulz's construction rules shown in table 1.5. OMPs are assembled so that the proteins hydrophobic surface is exposed to the lipid bilayer core. In simple terms a beta barrel protein consists of a large beta-sheet that twists to form an enclosed assembly where the first strand is hydrogen bonded to the last. These proteins are often characterised by short membrane spanning strands of alternate hydrophobic residues up to ten amino acids in length, followed by loop regions which are often highly variable (Schulz, 2000).

OMPs across many Gram-negative bacteria fulfil multiple roles from structural integrity (Pautsch & Schulz, 1998) and adhesion (Vandeputte-Rutten et al., 2003, Yamashita et al., 2011, Prince et al., 2002) to roles as porins (Hong et al., 2006, Oomen et al., 2004, Koronakis et al., 2000, Cowan et al., 1992) and enzymes (Snijder et al., 1999, Hwang et al., 2002). Considering that 2-3% of genes found within Gram-negative bacteria are predicted to encode barrel like OMPs only around ~25 of these structures have been solved to date leaving a huge gap in our

knowledge of this important protein family. Of the solved structures (shown in table 1.6) many appear as monomers, one from *Escherichia coli* or *E. coli* (OmpLA) has been demonstrated to form a homodimer (Snijder et al., 1999) and several porins from *E. coli* (tolC, OmpF, PhoE) have shown homo-trimer structures (Koronakis et al., 2000, Cowan et al., 1992). On the whole these proteins are characterised with long flexible loops extending at the extracellular side in contrast with the short tight turns extending from the periplasmic side of the membrane. The beta barrel structure can be described like inverted micelles where amino acid residues on the surface of the barrel which are facing the lipid bilayer are exclusively hydrophobic residues with an internal core of mostly polar amino acids. This is clearly the opposite of what is seen with most soluble proteins which harbour a hydrophobic core.

I.	The N and C termini are both located at the periplasmic side and the number of β-strands is even*
II.	The β-strands tilt is always 45^{o} are comparable to the β-sheet twist.
III.	All of the β-strands are antiparallel and connected to their closest neighbour along the chain.
IV.	The shear number of an n-stranded barrel is positive and around n+2.
V.	The connection strands at the periplasmic ends are short turns of only a few residues and are designated T1, T2, etc.
VI.	The external barrel loops are normally long and named L1, L2, etc.
VII.	The β-barrel surface is in contact with the non-polar membrane interior and contains aliphatic side chains which form a non-polar ribbon.
VIII.	The aliphatic ribbon is lined by two girdles composed of aromatic side chains.
IX.	The sequence of all parts of the β-barrel is variable during evolution when compared to soluble proteins with the external loops showing exceptional sequence variability.

Table 1.5. The principles of a β-barrel fold. A summary of the β-barrel construction rules adapted from Schulz, 2000. *An exception to the even strand rule is seen in the 19-stranded voltage-dependent anion channel (VDAC) within mitochondria (Maurya & Mahalakshmi, 2013) and chloroplasts (Hayat & Elofsson, 2012).

Protein	No. β-strands	Oligomeric state	Organism	No. Amino acids	PDB code	Function	Reference
OmpA	8	Monomer	*E. coli*	171	1QJP	Structural, adhesion, invasion	Pautsch & Schulz, 1998
OmpW	8	Monomer	*E. coli*	197	2F1V	Hydrophobic channel	Hong et al., 2006
OmpX	8	Monomer	*E. coli*	148	1QJ8	Toxin binding	Vogt & Schulz, 1999
PagP*	8	Monomer	*E. coli*	170	1MM4	Palmitoyl transferase	Hwang et al., 2002
NspA	8	Monomer	*N. meningitides*	155	1P4T	Adhesion	Vandeputte-Rutten et al., 2003
Ail	8	Monomer	*Y. pestis*	157	3QRC	Adhesion	Yamashita et al., 2011
OmpT	10	Monomer	*E. coli*	297	1I78	Protease	Vandeputte-Rutten et al., 2001
OpcA	10	Monomer	*N. meningitides*	253	1K24	Adhesion	Prince et al., 2002
NalP	12	Monomer	*N. meningitides*	308	1UYN	Auto-transporter	Oomen et al., 2004
OmpLA	12	Dimer	*E. coli*	269	1QD6	Phospho-lipase	Snijder et al., 1999
TolC	3X4	Trimer	*E. coli*	428	1EK9	Export channel	Koronakis et al., 2000
FadL	14	Monomer	*E. coli*	427	1T16	Fatty acid transporter	Van den Berg et al., 2004
OmpF	16	Trimer	*E. coli*	340	2OMF	Porin	Cowan et al., 1992
PhoE	16	Trimer	*E. coli*	330	1PHO	Porin	Cowan et al., 1992

Porin	16	Trimer	*Rh. capsulatus*	301	2POR	Porin	Weiss & Schulz, 1992
Porin	16	Trimer	*Rh. blastica*	289	1PRN	Porin	Kreusch et al., 1994
OmpK 36	16	Trimer	*K. pneumoniae*	342	1OSM	Porin	Dutzler et al., 1999
Omp32	16	Trimer	*C. acidovorans*	340	1E54	Porin	Zeth et al., 2000
MspA	8X2	Octamer	*M. smegmatis*	184	1UUN	Porin	Faller et al., 2004
LamB	18	Trimer	*E. coli*	421	1MAL	Maltose porin	Schirmer et al., 1995
Maltop orin	18	Trimer	*S. typhimurium*	421	2MPR	Maltose porin	Meyer et al., 1997
ScrY	18	Trimer	*S. typhimurium*	413	1A0S	Sucrose porin	Forst et al., 1998
BtuB	22	Monomer	*E. coli*	594	1NQE	Cobalamin transporter	Chimento et al., 2003
FhuA	22	Monomer	*E. coli*	723	2FCP	Iron transporter	Ferguson et al., 1998
FepA	22	Monomer	*E. coli*	724	1FEP	Iron transporter	Buchanan et al., 1999
FecA	22	Monomer	*E. coli*	741	1KMO	Iron transporter	Ferguson et al., 2002

Table 1.6. β-barrel outer membrane proteins with solved structures. A summary of outer membrane proteins with solved structures. *Structure solved by NMR.

1.4.3.2 *Escherichia coli* - OmpA

Outer membrane protein A (OmpA) is one of the most studied and characterised proteins within the family. The *ompA* gene, originally annotated as tolG was first identified in *E. coli* in 1974 and although the study did not identify the protein as a major OMP they did conclude that a mutation within the *ompA* gene gave rise to a phenotype which was tolerant to a L-type colicin known as Bacteriocidin JF246 (Foulds & Chai, 1978). Shortly after its identification the same pair demonstrated that the protein was indeed a major outer membrane protein with a later study adding that it must be comprised of two domains as C-terminal truncated mutants could still insert themselves into the membrane with an expected phenotype (Bremer et al., 1982). Following this early work, Morona et al. postulated that the protein may cross the membrane several times following interpretation of colicin resistance patterns developed from ompA mutants (Morona et al., 1984). Two years later an extremely elegant study based on Raman spectroscopy by Vogel & Jähnig concluded that the OmpA protein was composed of eight membrane spanning anti-parallel β-strands connected by four long outer loops which are exposed at the surface of the membrane with smaller or short turns on the periplasmic side (Vogel & Jähnig, 1986). Not until 1998 was Vogel & Jähnig's prediction confirmed by the release of the first single crystal OmpA N-terminal structure (Pautsch & Schulz, 1998), their prediction was remarkably accurate and the crystal structure (figure 1.11) showed that the N-terminal 171 amino acids form an eight stranded β-barrel often described as an inverted micelle with a hydrophobic girdle and a predominantly polar core. The C-terminal periplasmic domain is responsible for binding to peptidoglycan. The early exploration of integral OMPs led to the creation of a set of β-fold construction rules described in a review by Schulz (Schulz, 2000). These principles are summarised in table 1.5.

The heat modifiable protein OmpA from *E. coli* was originally purified in 1977 and was demonstrated to have a molecular mass of 33 kDa. This value however was seen to differ according to experimental conditions. When the whole protein was subjected to 60^{o}C heat treatment, OmpA was observed to migrate to 35-36 kDa (Beher et al., 1980, Freudl et al., 1986) in contrast when subjected to SDS-PAGE without heat treatment the protein is seen to resolve at around 28-31 kDa (Freudl et al., 1986). This difference in migration is likely due to the differences in conformation or denaturation of the heat treated sample compared to the native sample, this difference in unfolding allows a larger amount of SDS to bind to the protein resulting in the varying migration.

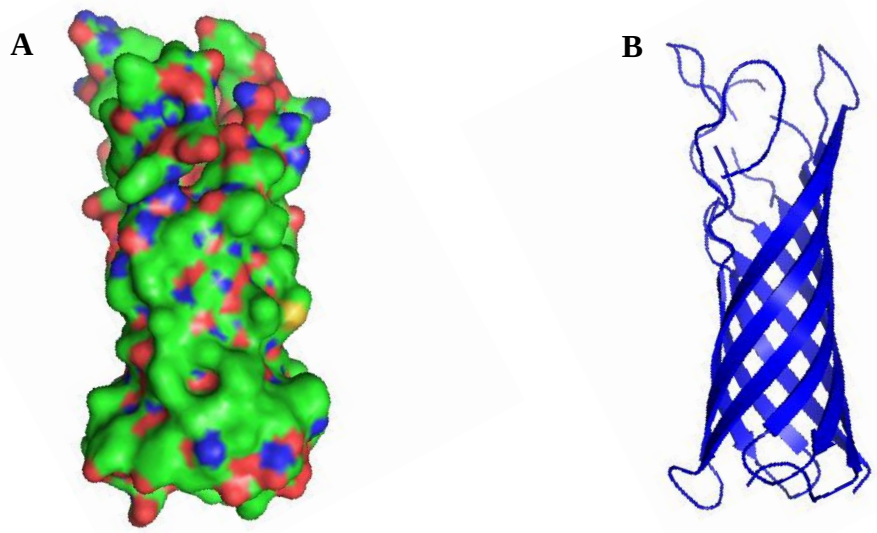


Figure 1.11. The crystal structure of the *E. coli* OmpA membrane spanning domain. A. The surface representation of OmpA. B. A cartoon of the secondary structure components of OmpA. Both figures were generated using PDB accession code 1BXW (Pautsch and Schulz, 1998) and images produced in Pymol (Schrödinger, 2010).

OmpA from *E. coli* has been implicated in various cellular and pathogenic roles including structural integrity and cell shape (Prasadarao et al., 1996), a receptor of colicin (Foulds & Barrett, 1973) a receptor for several phages (Morona et al., 1984), roles in F-conjugation (Weiser & Gotschlich, 1991) and invasion of brain microvascular endothelial cells or BMEC's (Shin et al., 2005).

1.4.3.3 *E. coli* – OmpX

OmpX from *E. coli* is an 18 kDa integral membrane protein which has been implicated in cell adhesion and invasion of host cells (Vogt & Schulz, 1998, de Kort et al., 1994) and the evasion of the host complement system (Vogt & Schulz, 1999). When comparing the X-ray crystal structure (figure 1.12) with its counterpart OmpA, it becomes apparent that there is a notable difference in the global dimensions and OmpA contains significantly longer extracellular loops. Curiously it has been suggested that a portion of the OmpX β-barrel protrudes out of the exterior with this exposed edge of β-sheet able to bind many proteins that have a surface exposed β-strand (Schulz, 2000). As OmpX is often highly expressed during *E. coli* stress

conditions, OmpX may be a useful defensive tool by binding attacking proteins (Schulz, 2000). Alike OmpA both barrels are packed with internal polar residues forming a hydrogen bonding network (Pautsch & Schulz, 1998, 2000, Vogt & Schulz, 1999) and as they can be described as inversed micelles they are unlikely to form pores (Schulz, 2000).

Alike OmpA, various OmpX homologs have been identified among Gram-negative bacteria, these include the attachment invasion locus (Ail) protein from *Yersinia pestis* (Yamashita, 2011), the phoP activated gene C protein (PagC) (Nishio et al., 2005) and the resistance to complement killing protein (Rck) (Heffernan et al., 1992) proteins from *Salmonella typhimurium*, which are further discussed in section 1.4.3.5.

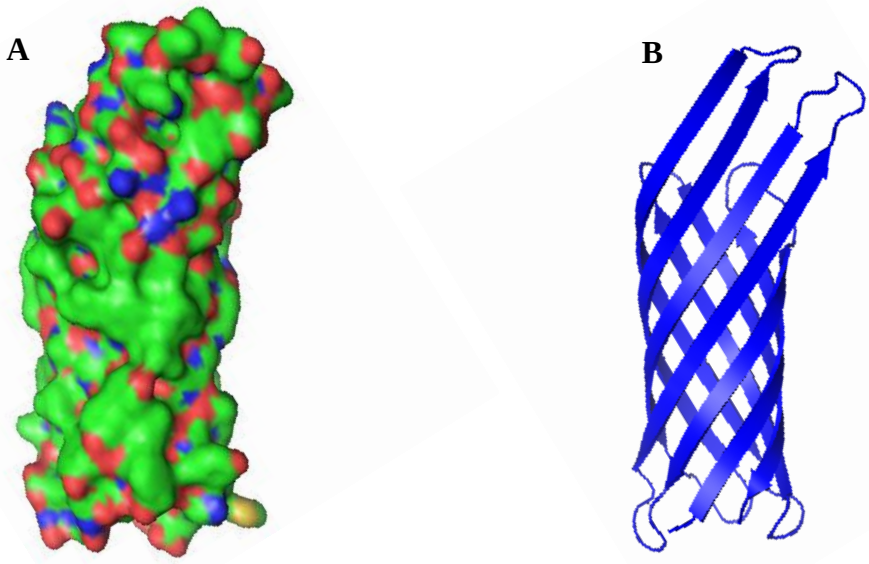


Figure 1.12. The crystal structure of *E. coli* OmpX. **A**. The surface representation of OmpX. **B**. A cartoon of the secondary structure components of OmpX. Both figures were generated using PDB accession code 1QJ8 (Vogt & Schulz, 1999) and images produced in Pymol (Schrödinger, 2010).

1.4.3.4 *E. coli* – Omp W

OmpW also from *E. coli* represents another well conserved integral membrane protein. The 21 kDa β-barrel protein (figure 1.13) spans the outer membrane of *E. coli* and has been demonstrated to behave as a hydrophobic channel (Hong et al., 2006). Several homologs have been identified across Gram negative bacteria and implicated in environmental stress mechanisms. These roles include protection from oxidation (Gil et al., 2007), osmosis (Xu et

al., 2005) and temperature (Nandi et al., 2005). Proteomic studies have also identified some correlation between the upregulation of OmpW and bacterial antibiotic resistance and increased virulence (Goel and Jiang, 2009). More recent work has demonstrated that *ompW*-deletion mutants are more susceptible to phagocytosis and the protein is an efficient protective antigen (Wu et al., 2013). However, this may not be of clinical importance due to the association of increased OmpW expression with reactive arthritis (Singh et al., 2007).

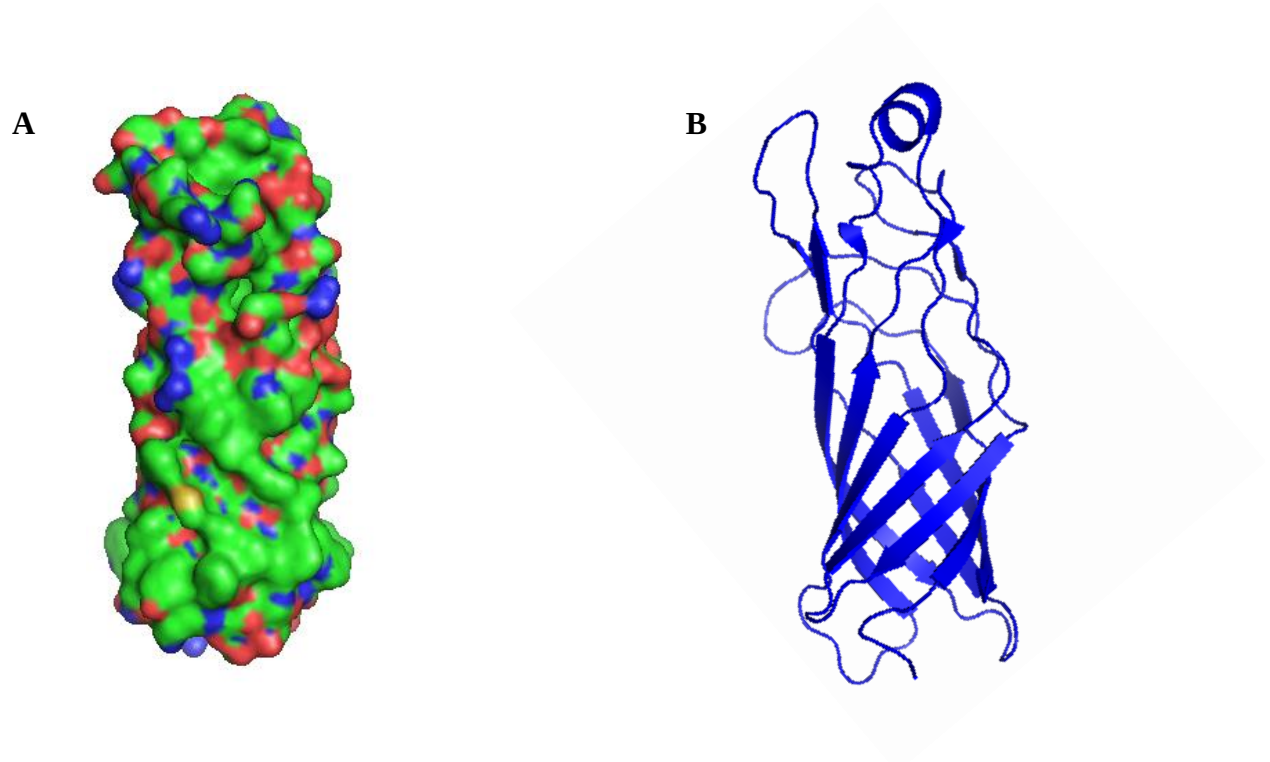


Figure 1.13. The NMR determined structure of *E. coli* OmpW. A. The surface representation of OmpW. B. A cartoon of the secondary structure components of OmpA. Both figures were generated using PDB accession code 2F1V (Hong et al., 2006) and images produced in Pymol (Schrödinger, 2010).

1.4.3.5 OmpA-like proteins

OmpA-like homologues can be found across many Gram-negative and positive bacteria and are designated as the OmpA-like superfamily. This family contains a whole host of proteins with varying levels of characterisation. The protein OmpA from *E. coli* is well studied yet others lack any functional annotation. Several examples of characterised OmpA-like proteins are further explained in table 1.7. As many bacteria possess a high number of these OmpA-like proteins which are surface exposed, within pathogenic strains they are often implicated in a

number of pathogenic roles in various organs but are most commonly associated with the urogenital and respiratory system (Krishnan and Prasadarao, 2012). Their involvement in disease is often attributed to roles including adhesion, evasion, invasion and host cell activation (Smith et al., 2007) and with their surface exposed position alongside a high copy number within the bacterial outer membrane they can be promising targets for vaccine development, with some examples shown in table 1.8.

Species	Designated protein name	Organ affected	Known pathogenic roles	Reference
Actinobacillus	OmpA	Respiratory tract	Adherence	Ojha et al., 2005
Aggregatibacter	OmpA	Oral	Adherence	Kajiya et al., 2011
Chlamydia	MOMP	Urogenital	Adherence	Su et al., 1996
Haemophillus	P5	Respiratory tract and CNS	Adherence and evasion	Miyamoto & Bakaletz, 1996
Klebsiella	OmpA	Respiratory tract	Adherence, proinflammatory and antimicrobial resistance	Catalina et al., 2011 Llobet et al., 2009
Leptospira	Loa22	Nephritis	Toxicity	Ristow et al., 2007
Neisseria	OmpA/Rmp	Genitourinary	Adherence, invasion and evasion	Serino et al., 2007
Salmonella	OmpA	Systemic	Unknown	Stocker et al., 1979
Yersinia	OmpA	Systemic	Intracellular survival	Bartra et al., 2011

Table 1.7. A summary of characterised OmpA homologues across various bacterial species.

Bacteria	Protein	Evidence	Reference
E. coli	OmpA	Activates dendritic cells	Jang et al., 2011
Haemophilus influenzae	OmpP1	Protective in animal models	Bolduc et al., 2000
Haemophilus influenzae	OmpP2	Conserved across all species	Galdiero et al., 2003
Klebsiella pneumonia	OmpA	Activates macrophages and dendritic cells	Catalina et al., 2011
Leptospira	LipL32	Reactivity with convalescent mouse sera and well conserved	Grassmann et al., 2012
Neisseria meningitidis	NspA	Antibodies are protective	Martin et al., 2000
Richettsia	OmpA	Protective in mice	Jiao et al., 2005
Shigella flexneria	OmpA	Well conserved and evokes type I cell mediated immunity	Pore & Chakrabarti, 2013

Table 1.8. A summary of proposed outer membrane protein vaccine targets.

1.4.3.5.1 *Yersinia pestis* – Ail

The OmpA homolog, Ail is spread throughout several species of pathogenic *Yersiniae* and is of particular interest within the plague pathogen *Y. pestis*. It is a well-studied, multifunctional protein and a member of the OmpA-like family. Within pathogenic *Yersiniae*, Ail has five known functions,

i. Complement inhibition/evasion (Kolodziejek et al., 2007)
ii. Inhibition of the inflammatory response (Felek and Krukonis, 2009, Hinnebusch et al., 2011)
iii. Adhesion and attachment to host tissues (Kolodziejek et al., 2010, Felek & Krukionis 2009)
iv. Delivery of *Yersinia* outer proteins to host tissues (Felek & Krukonis, 2008, Yamashita et al., 2011)
v. Prevention of the recruitment of polymorphonuclear leukocytes to the lymph nodes

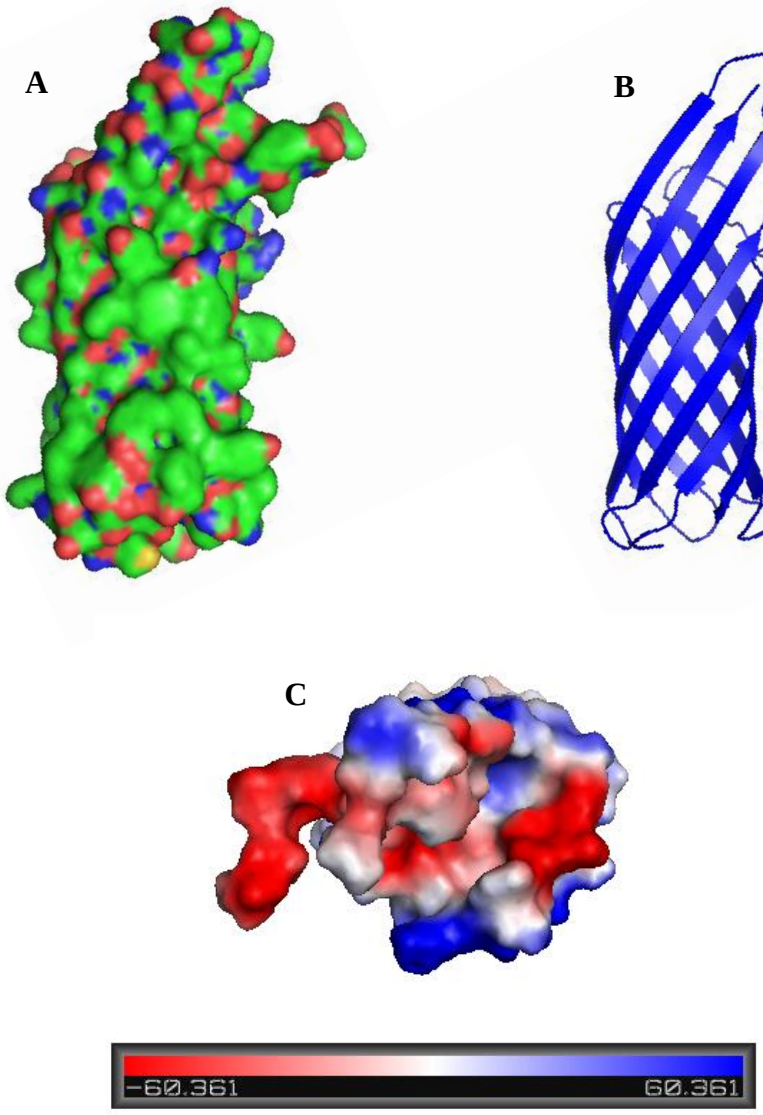


Figure 1.14. The crystal structure of *Y. pestis* Ail. **A**. The surface representation of Ail. **B**. A cartoon of the secondary structure components of Ail. **C**. Surface representation of the electrostatic potential of Ail. The figure is generated looking through the barrel from the exterior of the membrane. All figures were generated using PDB accession code 3QRA (Yamashita et al., 2011) and images produced in Pymol (Schrödinger, 2010).

Structurally Ail is a relatively small, monomeric, 8-stranded β-barrel at 17 kDa. The crystal structure as shown in figure 1.14 is typical of the OmpA-like family with 8 antiparallel β-strands and four extracellular loops. Alike OmpA and OmpX there appears to be no apparent channel throughout the transmembrane barrel (Yamashita et al., 2011) although the two regions

of positive charge around the extracellular surface may play a role in Ail's ability to bind a number of substrates (Yamashita et al., 2011).

1.4.3.5.2 *Neisseria* – NspA

The 18 kDa, 8-stranded β-barrel, membrane protein NspA is remarkably well conserved among *Neisseria gonorrhoeae* and *N. meningitides* isolates (Martin et al., 1997, Plante et al., 1999, Moe et al., 1999) however, its function is poorly understood. The crystal structure was published during 2003 as shown in figure 1.15 (Vandeputte-Rutten et al., 2003) and indicated that one extracellular loop (L3) was particularly large and of possible therapeutic value due to its surface exposure *in vivo*. Searches based on sequence similarity indicate that NspA holds significant sequence similarity to the Opa family of *Neisseria* OMP proteins which are understood to behave as adhesins (Martin et al., 1997) highlighting a role in virulence.

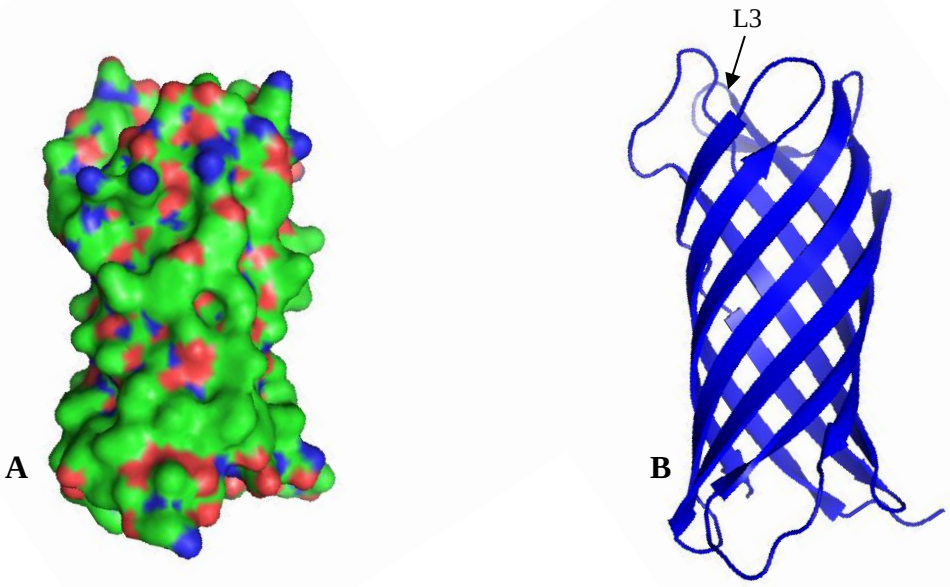


Figure 1.15. The crystal structure of *Neisseria meningitidis* NspA. **A**. The surface representation of NspA. **B**. A cartoon of the secondary structure components of NspA. Loop L3 is exposed *in vivo* and of possible therapeutic value (Vandeputte-Rutten et al., 2003). Both figures were generated using PDB accession code 1P4T (Vandeputte-Rutten et al., 2003) and images produced in Pymol (Schrödinger, 2010).

1.4.3.6 OmpA-like proteins and the blood brain barrier

The blood-brain barrier (BBB) is an essential component of the neural microenvironment and is formed by brain microvascular endothelial cells (BMECs), astrocytes and pericytes which create a physical yet functional barrier that regulates the passage of molecules to and from the brain (Kim, 2006). The barrier itself protects not only from normal fluctuations and changes within the host but also safeguards the brain from any microorganisms and toxins that have gained access to the circulatory system. Pathogens have been described to cross the BBB in three distinct ways, these include transcellular and paracellular movement across the barrier or gaining access by means of the 'Trojan-horse' mechanism as described in figure 1.16.

Transcellular movement by bacteria across the BBB refers to the transfer of pathogens through the actual cell with no disruption to the tight junctions (Kim, 2006). In contrast to this, paracellular transfer is the crossing of pathogens between the barrier cells sometimes with some form of tight junction disruption (Kim, 2006). The third means of crossing involves the Trojan-horse mechanism whereby the pathogen gains access to the neural microenvironment by migrating across the barrier inside of phagocytes, for example within an infected macrophage (Kim, 2006).

Infections of the central nervous system remain an important health concern with bacterial meningitis within the top ten causes of infection-related deaths worldwide (World Health Organisation, 2004). For most meningitis causing pathogens the route of BBB transfer is most often transcellular with evidence to support this type of crossing seen within *E. coli, Streptococcus agalactiae, S. pneumoniae, Neisseria meningitides* and the fungal pathogen *Cryptococcus neoformans*. Several studies have suggested that the protozoan pathogen Trypanosoma spp (Masocha et al., 2007, Grab et al., 2004) and some strains from *Borrelia* spp (Grab et al., 2005, Sethi et al., 2006) may follow a paracellular route when crossing the blood-brain barrier. However, this still remains poorly understood and they have also been shown able to cross a BMEC barrier via a transcellular mechanism (Grab et al., 2005, Sethi et al., 2006). Both *Listeria monocytogenes* and Mycobacterium *tuberculosis* have been suggested to utilise the Trojan-horse mechanisms of transfer whereby they gain access to the neural environment via infected phagocytes. However, alike those pathogens implicated in paracellular transfer both *L. monocytogenes* and *M. tuberculosis* are able to use a transcellular route.

Within some infections there appears to be a relationship between the level of bacteraemia and the development of meningitis particularly during infection with *E. coli*, *S. agalactiae* and *S. pneumoniae* (Kim et al., 1992, Ferrieri et al., 1980, Kim, 1987, Dietzman et al., 1974, Bell et al., 1985, Sullivan et al., 1982, Marra & Brigham, 2001). Research suggests that a certain degree of bacteraemia must be reached before the BBB is breached (Kim et al., 1992, Ferrieri et al., 1980, Kim, 1987, Dietzman et al., 1974, Bell et al., 1985, Sullivan et al., 1982, Marra & Brigham, 2001). This is particularly interesting as *Borrelia* spp are able to cross the BBB but bacteraemia is not normally high during the spirochaetal infection (Eshoo et al., 2012). Although during *E. coli* infections a high level of bacteraemia is required for BBB invasion this state is not sufficient for the development of meningitis and the actual binding of HBMECs is a requirement for the bacterial *in vivo* crossing of the BBB (Kim, 2006).

A key example of this blood brain barrier crossing involves the OmpA protein from certain pathogenic strains of *E. coli*. These strains have the ability to trigger septicaemia and meningitis in both humans and domestic animals (Smith et al., 2007). A key study indicated that purified OmpA protein from these strains can adhere to brain microvascular endothelial cells or BMECs (Shin et al., 2005). The same study highlighted that a pre-incubation of BMEC's with the beta barrel region only, (i.e. the loops deleted), failed to reduce adhesion underlining the importance of OmpA within adherence and evasion (Prasadarao et al., 1996). Isogenic *E. coli* K1 mutants with gene deletions corresponding to the sequence encoding OmpA which is known to contribute to BMEC binding were significantly less able to cross the blood-brain barrier than the wildtype counterpart (Kim, 2001).

Outside of the meningitic *E. coli* strains a similar pathway can be observed *within Enterobacter sakazakii*, a food-borne pathogen which can lead to sepsis and meningitis within paediatric patients (Chenu & Cox, 2009). The OmpA homolog within *E. sakazakii* is again crucial for adherence and the invasion of BMEC's and human enterocytes (Kim et al., 2010, Wang & Kim, 2002).

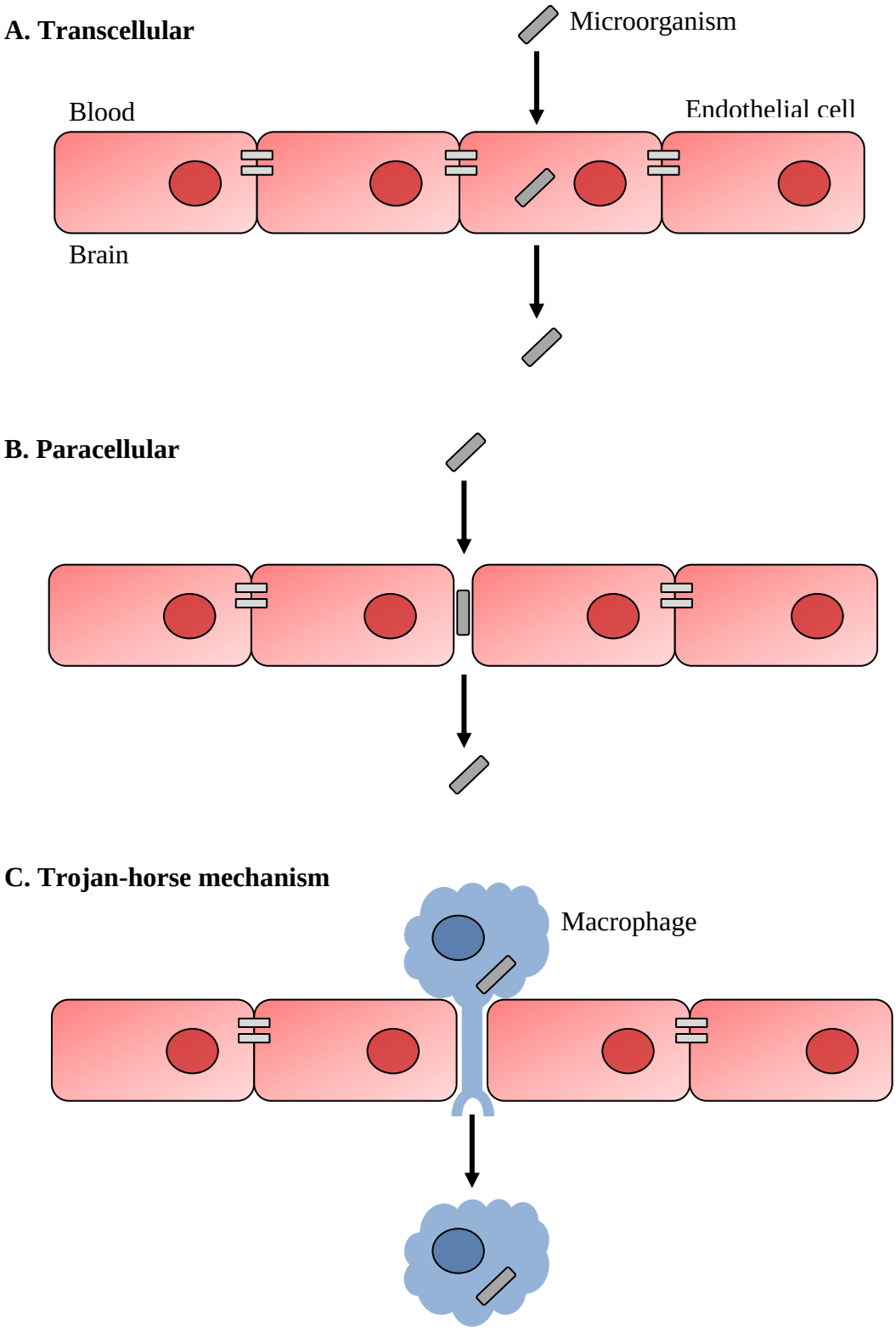


Figure 1.16. Mechanisms of microbial traversal of the blood-brain barrier. The blood brain barrier is a structural and functional barrier which regulates the passage of molecules both in and out of the protected environment. The barrier itself is formed by brain microvascular endothelial cells (BMECs), astrocytes and pericytes. The arrangement of BMECs allows the formation of tight junctions which restrict the passage of molecules into the neural environment (continued…).

The three major routes across the blood-brain barrier made by microbes. **A**. Transcellular – Whereby the pathogen crosses the barrier through barrier cells with no evidence of tight junction access or disruption. **B**. Paracellular – Pathogens are observed to cross the barrier by disrupting the tight junctions or can be found between barrier cells. **C**. Trojan-horse mechanism – Pathogens gain access via infected phagocytes.

1.4.4 Known Borrrelia OMPs

As previously noted very few *B. burgdorferi* integral membrane proteins have been identified and many of these remain poorly understood. Within all bacteria OMPs play a wide variety of important roles including nutrient acquisition via porins, antibiotic resistance by means of efflux pumps and adhesion (Schulz, 2002). As freeze fracture electron microscopy identified that the *Borrelia* outer membrane possesses 10 fold fewer integral OMPs in contrast to the *E. coli* outer membrane the identification of these important membrane components has been hindered.

Very few porins are known within *B. burgdorferi* but include P13, Oms28, DipA, BesC and P66 within relapsing fever *Borreliae*. Both P13 and Oms28 remain poorly characterised but are speculated to hold porin like properties. BesC has been described as part of the efflux system (Bunikis et al., 2008) and DipA with dicarboxylate specificity (Thein et al., 2012).

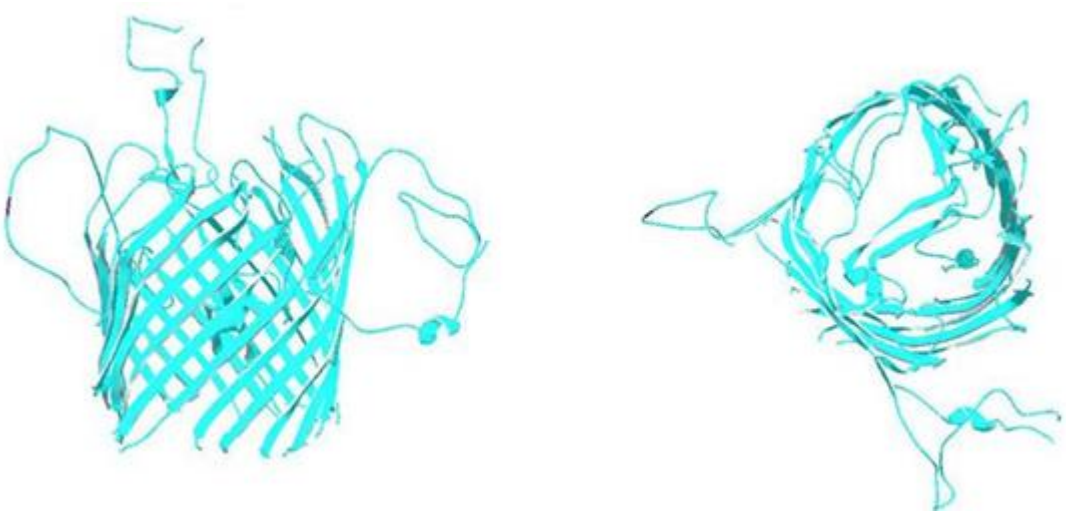


Figure 1.17. P66 homology modelling. The mature P66 amino acid sequence from *B. burgdorferi* was processed by the PRED-TMBB sever and the tertiary structure predicted. The protein is displayed as a ribbon model. Figure taken from Kenedy et al., 2014.

DipA also known as Oms38 was originally identified in *Borrelia* strains responsible for relapsing fever and was extracted from OM fractions of *B. duttonii*, *B. hermsii* and *B. recurrentis*. These extractions were also demonstrated to hold pore forming activity *in vitro* (Thein et al., 2012). Further investigations found that DipA was well conserved among relapsing fever strains (Thein et al., 2012) and a homolog with similar characteristics was noted among Lyme disease *Borreliae* known as BB0418. Although the structure is yet to be solved

the DipA protein is believed to be an integral outer membrane porin and is fairly conserved across *Borrelia* strains.

P66 however is the best characterised, non-selective porin, which also displays adhesion properties (Coburn & Cugini, 2003). P66 which is encoded by the open reading frame *bb0603* has a calculated molecular weight of 66kDa and is predicted to form a β-barrel comprised of a 22 or 24 stranded transmembrane domain with modelling data shown in figure 1.17. Phase partitioning and circular dichroism studies support the hypothesis and suggest the protein is amphiphilic and consists of 48% β-strand (Kenedy et al., 2013). *In vitro* studies have demonstrated that not only can P66 form pores in artificial membranes with an extremely high channel conductance (Lugtenberg and van Alphen, 1983). It can also bind to β3-integrin and due to its surface exposed domains the protein is of particular interest for vaccination development (Kenedy et al., 2013). Unlike other known bacterial OMPs P66 does not display heat modification during heat shift assays and it has been postulated that this may be a unique property of spirochetal OMPs (Kenedy et al., 2013, Dyer et al., Appendix 4).

1.5 Protein expression throughout the enzootic lifecycle

1.5.1 Major surface proteins of *Borrelia* expressed within the tick vector

Following attachment to a rodent host, *Borrelia* can be identified within the tick midgut within 24 hours of the blood meal (Schwan & Piesman, 2000). During this time the number of *Borrelia* within the tick midgut remains fairly stable (Dunham-Ems et al., 2009) until blood influx occurs which triggers essential environmental changes. This blood influx is characterised by an increased environmental temperature, a reduced pH and a rapid change in the number of *Borrelia* residing within the midgut upto one hundred thousand *Borrelia* cells per tick (Pulzova & Bhide, 2014). This ultimately leads to a swift modification to the *Borrelia* transcriptome and importantly the *Borrelia* outer membrane (Pulzova & Bhide, 2014). These global changes are actioned by the response regulator, Rrp2-RpoN-RpoS pathway which is responsible for the induction of more than 10% of *Borrelia* genes (Xu et al., 2010). The initial feeding induces expression of the outer membrane protein p66 (Cugini et al., 2003) and a family of proteins known for their ability to evade the complement response known as, complement regulator-acquiring surface proteins or CRASPs (Pulzova & Bhide, 2014). These early changes to the *Borrelia* surface is likely to protect the bacteria from destruction during the primary stages of the blood meal (Pulzova & Bhide, 2014).

During the later stages of the blood meal the protein BBO250 appears crucial for both correct cell division and cell envelope integrity (Liang et al., 2010). As the *Borrelia* count within the midgut begins to rise BBO250 is induced. Alongside BBO250 the outer surface protein A or OspA is abundantly expressed during the tick midgut stages (Schwan & Piesman, 2000). Although OspA has multiple functions (Pulzova & Bhide, 2014), during the tick midgut stage it is predominantly responsible for the mediation of attachment to the tick endothelium via the tick receptor for OspA (TROSPA) resulting in tick gut colonisation (Pulzova & Bhide, 2014). Other studies have also suggested that OspA may also act as its own receptor increasing adherence by bacterial aggregation (Pulzova & Bhide, 2014).

The 28 kDa lipoprotein OspA is highly immunogenic and is anchored to the *Borrelia* outer membrane by a lipidated N-terminal cysteine (Huang et al., 1998). It forms a twenty one stranded antiparallel β-sheet with the C-terminus containing a single α-helix as shown in figure 1.18. The monomeric protein is composed of two globular domains at each side of the β-sheet which is connected by the non-globular middle section (Bu et al., 1998). Both SAXS and NMR

studies suggest the protein exists in solution as a monomer (Bu et al., 1998, Huang et al., 2001) and therefore does not have a typical hydrophobic core.

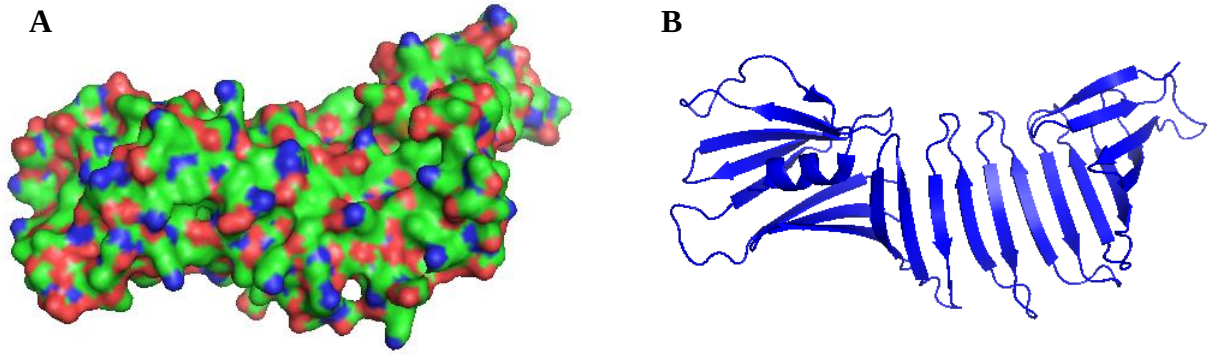


Figure 1.18. The crystal structure of *B. burgdorferi* OspA. A. The surface representation of OspA. **B**. A cartoon of the secondary structure components of OspA. Both figures were generated using PDB accession code 2G8C (Makabe et al., 2006) and images produced in Pymol (Schrödinger, 2010).

Although both OspA and OspB are expressed upon *Borrelia's* surface when residing within an unfed tick midgut, these proteins are downregulated upon the acquisition of a blood meal. During the feeding and following a rapid increase of the *Borrelia* cell density within the tick midgut, the migration of *Borrelia* to the salivary gland begins. This migration is coordinated by a number of interactions between *Borrelia's* surface proteins and the ticks own components. The lipoprotein BBE31 has been demonstrated to interact with a tick gut receptor TRE31 and is required for *Borrelia's* translocation into the haemolymph (Kumar et al., 2011) and *Borrelia's* BBA52 surface protein has also been shown to assist the migration of the bacteria to the tick salivary glands (Kumar et al., 2011).

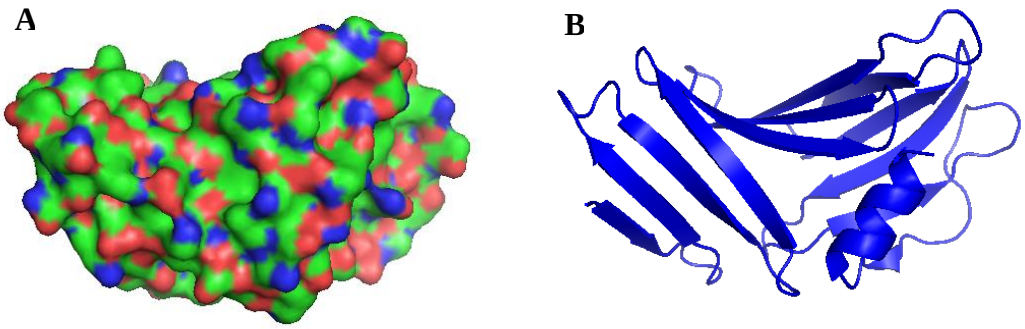


Figure 1.19. The crystal structure of OspB. A. The surface representation of OspB**. B**. A cartoon of the secondary structure components of OspB. Both figures were generated using the PDB accession code 1RJL (Becker et al., 2005) and images produces in Pymol (Schrödinger, 2010).

The most studied *Borrelia* protein within this stage is the outer surface protein C (OspC). Alongside BBA64 these proteins are not abundantly expressed by *Borrelia* residing in the gut of an unfed tick but are upregulated during a blood meal (Pulzova & Bhide, 2014). Studies have demonstrated that OspC is involved in salivary gland invasion and binds tightly to tick salivary gland tissue (Pal et al., 2004) and BBA64 is crucial for mammalian infection via tick bite transmission (Gilmore et al., 2007, Gilmore et al., 2008). Additionally OspC can bind the tick salivary gland protein, Salp15 enabling survival of antibody mediated killing during initial infection (Ramamoorthi et al., 2005). Structurally, OspC is a 23 kDa lipoprotein composed of four long anti-parallel α-helices and one short α-helix as shown in figure 1.20. Dimerisation of OspC produces an α-helical hydrophobic core (Kumaran et al., 2001) and the protein has been demonstrated to bind to a number of host ligands (Pal et al., 2004, Pulzova & Bhide, 2014) likely due to its strong negative electrostatic charge highlighting its importance in human disease.

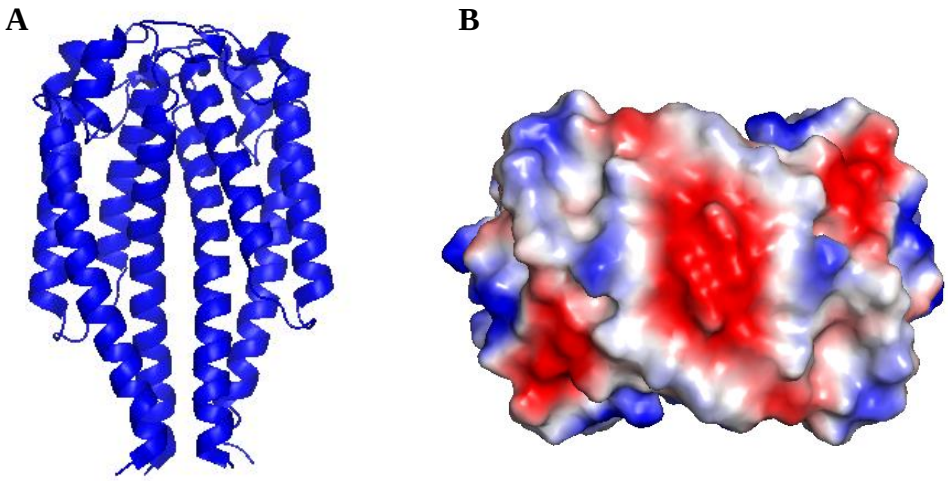


Figure 1.20. The crystal structure of OspC - A. Cartoon representation of the secondary structure components of OspC**. B**. The electrostatic potential of OspC viewed from the top of the molecule. Both figures were generated using PDB accession code 1GGQ with X-ray diffraction data from Kumaran et al., 2001 and images produced in Pymol (Schrödinger, 2010).

1.5.2 *Borrelia* proteins implicated in the early stage of host infection

During tick feeding, OspA and OspB are downregulated and OspC is more abundantly expressed on the *Borrelia* outer membrane. This shift continues through transmission and early infection which is reflected by the absence of OspA and OspB antibodies during early mammalian disease (Schwan and Piesman 2000). The down regulation of OspA at this time is particularly important. As previously mentioned OspA is highly immunogenic, it is a potent stimulator or neutrophils and the production of cytokines (Morrison et al., 1997) and it's down-regulation during initial infection is significant. However, OspA antibodies are detectable during chronic infection indicating that the protein plays a larger role (Hsieh et al., 2007). During these early stages of infection the skin acts as the initial hurdle. As the spirochaete penetrates this barrier via the tick bite, the next obstacle is to avoid immune recognition. The spirochaetes exposed lipoproteins are targeted by the innate immune system, where interactions are formed between pattern-recognition receptors and Toll-like receptors on the surface of dermal fibroblast, keratinocytes and Langerhans cells (Ebnet et al., 1997). An EM rash is often devoid of neutrophils as the invading *Borreliae* acquire host proteins upon their surface essentially shielding the pathogen from the host immune response (Szczepanski & Benach, 1991).

1.5.3 *Borrelia* proteins responsible for dissemination and extracellular matrix adhesion

For many pathogens the strategy to attach to host tissues is considered an essential component of colonisation and in general is mediated by numerous proteins expressed on a bacterial cell membrane known as adhesins. *In vitro* experiments have demonstrated that *Borrelia* can bind to a number of different host cells including endothelial cells, epithelial cells, lymphocytes, neuroglia and platelets (Comstock & Thomas, 1989, Szczepanski et al., 1990, Dorward et al., 1997, Garcia-Monco et al., 1989, Thomas & Comstock, 1989). Although numerous *Borrelia* surface proteins have been demonstrated to mediate adhesion to various components of the extracellular matrix (Antonara et al., 2007) the mechanisms in which *Borrelia* penetrate the ECM and move to more distant tissues remains poorly understood. Following a short period of multiplication during localised infection lasting approximately 2 days, the spirochaetes begin to penetrate the walls of small arteries leading to dissemination and access to remote sites of connective tissue throughout the body where they are able to interact with the extracellular matrix. Several surface proteins from *Borrelia* have been identified to facilitate adhesion to the

ECM whether this be by adhesion to fibronectin, proteoglycans such as decorin or type I collagen.

Proteoglycans are highly glycosylated proteins and are composed of a core protein which is covalently linked to one or more GAG chain. These chains are liner, long disaccharide repeats which are highly sulphated and can be classed by differences in their structure and their sensitivity to cleavage by lyases. Classes of GAGs are summarised in figure 1.21 but include heparin, heparan sulphate, dermatan sulphate and chondroitin-6-sulphate with various bacterial membrane proteins targeting each of these components. GAG binding specificity is also seen to differ throughout *Borrelia* strains and the targeting of different GAGs appears to result in the invasion of different host cell types. For example *Borrelia* strains which bind heparan sulphate appear to invade endothelial cells where those that target dermatan sulphate mainly attach to glial cells *in vitro* studies (Leong et al., 1998). Alongside this phenomenon, *Borrelia* strains that can adhere to both heparan sulphate and dermatan sulphate can invade both cell types (Leong et al., 1998, Pulzova & Bhide, 2014). This GAG specificity is explained by the fact that *Borrelia* produces multiple surface proteins which act as GAG-binding adhesins (Leong et al., 1998, Pulzova & Bhide, 2014, Parveen et al., 2006, Fischer et al., 2003) and this profile of expressed adhesins is likely to be characteristic for any given strain.

One example of this is the glycosaminoglycan (GAG) binding protein Bgp (BB0588) which is secreted and found on the surface of the outer membrane (Parveen et al., 2006). The protein plays important roles in early infection by the means of binding heparin and agglutinating erythrocytes but also has roles in the inactivation of toxic metabolites within the host (Pulzova & Bhide, 2014). Bgp however is not constantly expressed during infection (Parveen et al., 2006) which has led to researchers questioning its importance. However, as many *Borrelia* surface proteins have been identified to hold GAG-binding properties it is likely that constant expression of Bgp is simply not always required with an array of similar proteins also being available (Leong et al., 1998, Pulzova & Bhide, 2014, Parveen et al., 2006, Fischer et al., 2003)
.

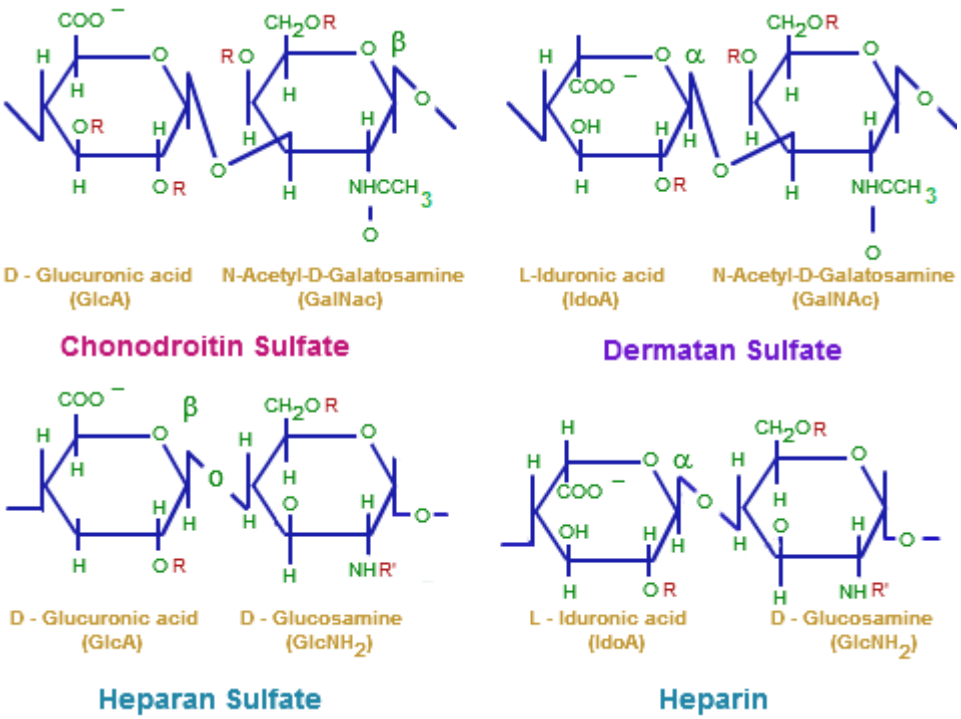


Figure 1.21. Glycosaminoglycan unit structure. The four common GAGs found within the ECM and their disaccharide structure. Chronodroitin sulphate is composed of the alternating sugars D-glucuronic acid and N-Acetyl-D-Galatosamine. Dermatan sulphate is composed of the alternating sugars L-lduronic acid and N-Acetyl-D-Galatosamine. Heparan sulphate is composed of D-Glucuronic acid and D-Glucosamine and the final GAG, heparin formed with the alternating sugars L-Iduronic acid and D-Glucosamine.

The variable surface protein 2 (Vsp2) produced by *B. turicatae* the relapsing fever spirochaete has also been implicated in the GAG binding strategy and can bind to the ECM components, heparin and chondroitin sulphate (Magoun et al., 2000). Structural studies on the lipoprotein provided insight into this interaction which appears to be governed by ionic interactions from the basic amino acids on the protein and the sulphate and carboxylate groups of the GAGs (Lawson et al., 2006). It was also observed that heparin binding sites on Vsp2 corresponded to shallow pockets of positive charge on the outer protein surface which agreed well with previous studies (Capila & Linhardt, 2002).

Other key proteins during the dissemination stages include the decorin binding proteins. These proteins are often implicated in the dissemination of *Borrelia* and the resulting arthritic symptoms (Fischer et al., 2003). Decorin binding protein A (DbpA – BBA24) and B (DpbB – BBA25) are the major players during this stage with distinct roles in cell binding (Fischer et al., 2003). A further method of adhesion lies within the binding to type I collagen, a protein

found within the connective tissue of the ECM. This binding has only been reported with the presence of decorin and results in the formation of protein fibres which enable the initial attachment leading to further invasion and the formation of micro-colonies of *Borrelia* within the ECM (Zambrano et al., 2004). Further mechanisms of adhesion include the porin P66 previously discussed in section 1.4.4. The P66 protein enables the spirochete to adhere to host cells via an interaction between itself and the integrins αIIβ3 and αvβ3 (Skare et al., 1997).

Throughout this stage the protein BBK32 is of particular interest and is localised on the cell surface of *Borrelia* during mammalian infection. The protein has the capacity to bind to fibronectin (Fischer et al., 2006) and interact with host GAGs show in figure 1.22 (Fischer et al., 2006). Studies suggest (Fischer et al., 2006) that not only does BBK32 promote binding to fibronectin but the protein is also involved in binding to purified heparin *in vitro*, dermatan sulphate and GAGs found on several cell lines including, 293 epithelial cells, C6 glial cells and EA-Hy926 endothelial cells (Fischer et al., 2006).

As the ECM is a known protected niche for pathogens, our understanding of how *B. burgdorferi s.l* gains access and persists despite an on-going humoral response to the spirochaetes surface antigens is essential. During this time the *Borrelial* proteome appears essential in order to extend to chronic infection (Kung et al., 2013).

1.5.4 Translocation of *Borrelia* across the vascular endothelium and the involvement of the blood brain barrier

Penetration of the vascular endothelium by *B. burgdorferi* remains confusing in terms of whether the pathogen follows a paracellular or transcellular route although most researchers support the theory that the spirochaete uses a transcellular route (Comstock et al., 1993). Experimental evidence has demonstrated that *B. burgdorferi* can translocate a human umbilical vein endothelial cell monolayer (HUVECs) across the cytoplasm (Comstock et al., 1993) and *B. burgdorferi* have also been found within the intercellular junctions and beneath monolayers of endothelial cells (Szczepanski et al., 1990) suggesting that spirochaetes may be able to pass between cells.

The actual penetration of the blood vessel wall consists of several processes including, dragging, tethering and adhesion interactions of the endothelium (Norman et al., 2008). These

interactions are mediated by the spirochaete cell surface binding to many host molecules including fibronectin, GAG's and integrins (Norman et al., 2008). More particularly dragging interactions involve the *Borrelia* protein BBK32 as shown in figure 1.22 and further discussed below.

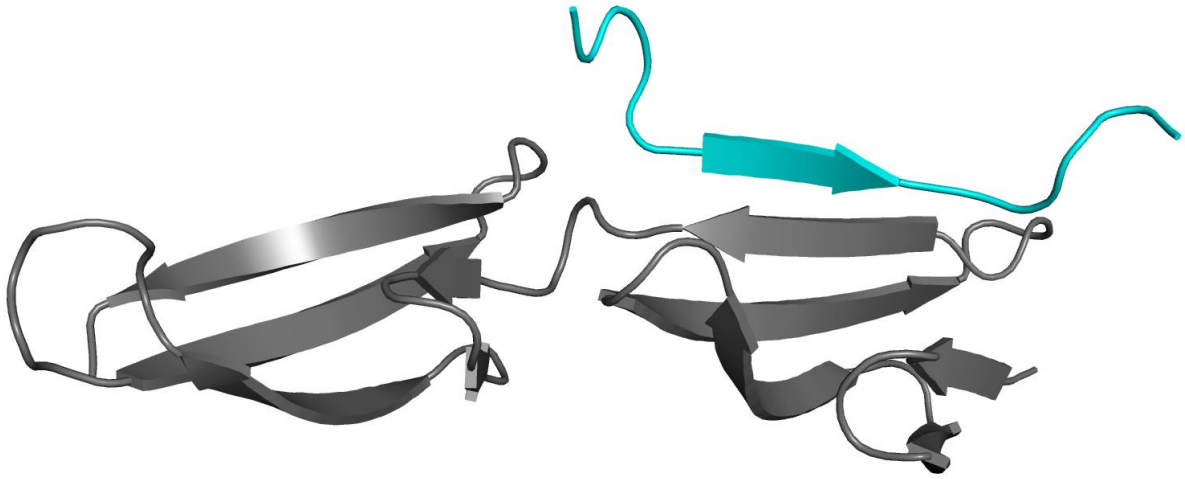


Figure 1.22. Crystal structure of BBK32 in complex with Fibronectin - Cartoon ribbon representation of the BBK32 fibronectin complex. GREY – Crystal structure of *B. burgdorferi* B31 BBK32. BLUE – Fibronectin molecule type 2-3 (FN^{2-3}).

Despite the expression of numerous surface proteins which can initiate a strong immune response, *B. burgdorferi* s.l is able to cross the blood-brain barrier and invade the CNS causing neuroborreliosis also providing the pathogen with a new immune privileged site (Berndtson, 2013). The mechanisms in which the pathogen crosses this barrier (Comstock et al., 1993, Szczepanski et al., 1990) and the damage that it causes (Berndston, 2013) which results in the neurological symptoms remain poorly understood. Viable *Borrelia* can be located within the CNS of Lyme borreliosis patients and viable spirochaetes have also have been observed intracellularly (Miklossy et al., 2008). Alike the ECM, *Borrelia* have been demonstrated to interact with many host cell surface molecules within both the blood-brain barrier and the CNS, these include decorin (Fischer et al., 2003), fibronectin (Fischer et al., 2006, Harris et al., 2014), GAGs (Magoun et al, 2000, Parveen et al., 2006, Leong et al., 1998) and integrins (Kenedy et al., 2013). These interactions with host molecules and *Borrelia's* ability to traverse the blood-brain barrier and invade the CNS appear to be facilitated by the numerous adhesion molecules

the spirochaete expresses upon its surface including, BBK32, Bgp, DbpA, DpbB and P66 whose combined affect is essential for this invasion (Pulzova and Bhide, 2014). In terms of adhesion numerous proteins from *Borrelia* are thought to contribute but more particularly stationary adhesion is thought to involve the *Borrelial* protein BBB07 which interacts with RGD-independent integrins including α3β1 which is expressed at intercellular junctions (Behera et al., 2008).

Some experimental data has demonstrated that *B. burgdorferi* OspA (*ospA*/BB_A15) alongside CD40 on the surface of brain-microvascular endothelial cells has a relationship which promotes tethering interactions and can be observed when the spirochaete crosses the blood-brain barrier (Pulzova et al., 2009). Further evidence for OspA's role in CNS invasion is supported by the fact that all *Borrelia* which adhere to neuronal cells are found to be OspA positive/OspC negative (Rupprecht et al., 2006) and the *ospA* gene has been described to be upregulated following the spirochaetes access to the cerebrospinal fluid (CSF) (Schutzer et al., 1997). This suggests that the OspA protein plays an important adhesion role during this time and is likely to be an essential factor to initiate Lyme Borreliosis (Schutzer et al., 1997).

1.6 Immune evasion strategies

Alike many other pathogens the spirochaete *Borrelia* employs numerous strategies in order to circumvent destruction by the hosts immune system. As recently reviewed by Berndston (2013) these strategies can be split into distinctive categories which include,

1. Exploitation of tick salivary proteins in order to prevent the initial immune response
2. The masking of *Borrelia* surface antigens to avert complement initiation.
3. The seizure of localised host plasminogen activation.
4. Continuous surface antigen variation.
5. The bypassing and dodging of immune molecules by both chemotactic and by reaching immune-privileged sites.
6. Horizontal gene transfer.
7. Quorum sensing and possible engagement in 'biofilm-like' behaviour.
8. Changes in the morphology of the spirochaete (formation of persister cells).

1.6.1 Tick salivary proteins

Throughout out the early stages of infection a number of tick proteins play important roles in enabling the transmitted bacteria to survive the barrage of host responses. A number of tick salivary proteins have anti-inflammatory, anti-haemostatic and immunomodulatory effects, these include ISL929 and ISL1373 which are involved in the downregulation of polymorphonuclear PMN integrin expression and have been implicated in the inhibition of superoxide production within the vicinity of the bite (Guo et al., 2009).

Salp20, another tick salivary protein has been shown to inhibit complement formation (Tyson et al., 2007). While a further interaction between *Borrelia* OspC and the tick salivary protein Salp15 ultimately blocks OspC-TLR2 interactions, effectively protecting the spirochaete from antibody mediated killing (Pulzova & Bhide, 2014).

Tick salivary proteins are also known to inhibit both the initial activation and the further proliferation of CD4+ T-cells by means of binding to the CD4 receptor region (Anguita et al., 2002). Furthermore these group of salivary proteins have also been demonstrated to inhibit numerous other immune cells including, dendritic cells (Hovius et al., 2008), natural killer cells (Kubes et al., 2002), macrophages (Gwakisa et al., 2001) and neutrophils (Montgomery et al., 2004). The contribution of these tick salivary proteins give the invading pathogen, *Borrelia* a modest advantage and in effect buy the spirochaete a little time to establish an infection before clearance by the host immune system (Pulzova and Bhide, 2014).

1.6.2 Hijacking plasminogen activation

Much like a number of other pathogenic bacteria, *B. burgdorferi* can hijack the host's plasminogen system in order to move beyond the initial site of infection (Brissette et al., 2009). In order to do so, the spirochaete must be able to cross the hosts physiological barriers such as the basement membrane and the extracellular matrix. This is often achieved by the recruitment of host proteases, normally the serine protease, plasmin. A number of pathogens express plasminogen receptors on their surface and the recruitment of host plasminogen provides the mechanism for invasion (Bhattacharya et al., 2012). As the inactive plasminogen is converted to active plasmin by host plasminogen activators the bacteria can now be described to carry a proteolytic surface which assists the pathogen to degrade and rearrange the ECM (Bhattacharya et al., 2012).

Curiously, plasminogen receptors have a high abundance of lysine residues so it is no surprise that the A-T rich genomes of *Borrelia* which gives rise to a high proportion of lysine residues within *Borrelia* proteins (Fraser et al., 1997) can act as molecules that bind plasminogen, granting not all of these may be of physiological importance. For example several plasminogen binding proteins are known from *Borrelia* including, ErpA/C/P (Brissette et al., 2009), OspA (Fuchs et al., 1994), OspC (Lagal et al., 2006) and a 70kDa protein (Hu et al., 1997). However some of these proteins are downregulated during mammalian infection suggesting that not all of these plasminogen binding proteins are essential during this stage of infection. During 2012 Önder et al identified that the *B. burgdorferi* outer surface protein C (OspC) contains a plasminogen receptor. Alike other pathogenic bacteria this surface receptor acts to recruit plasminogen which in turn is activated to plasmin allowing the digestion of fibrin and large glycoproteins essentially creating a path through the ECM (Önder et al., 2012). The relationship between OspC and invasiveness was further explored using mice. The study found that the invasiveness of *B. burgdorferi* using the mice model was dependent upon specific variations within the OspC protein (Lagal et al., 2006).

A further player in *Borrelia's* mechanism of hijacking host plasminogen is the cytosolic metalloenzyme, enolase (Toledo et al., 2011). Enolases also known as phosphopyruvate hydratases are responsible for the catalysis of 2-phospho-D-glycerate to phosphoenolpyruvate. Although enolases are often associated with the exterior of the cell, they are still considered cytosolic enzymes due to their lack of cell membrane anchoring and/or protein sorting machinery. These moonlighting cytosolic proteins are often identified as plasminogen receptors on the surface of both eukaryotic and prokaryotic cells (Pancholi, 2001, Pancholi & Fischetti, 1998, Bercic et al., 2008, LaFrentz et al., 2011, Mundodi et al., 2008, Pitarch et al., 2004, Pitarch et al., 2006).

1.6.3 Evading the complement system

The complement system is a complex component of innate immunity involving a highly regulated network of proteins and acts as the first line of defence against invading pathogens. (Sarma and Ward, 2011) During the late 19th century the German bacteriologist Hans Ernst August Buchner suggested that a 'principle' of blood serum was its ability to kill bacteria (Buchner et al., 1891). Shortly after, during 1896 a Belgian scientist working at the Pasteur Institute in Paris, Jules Bordet, demonstrated that the bactericidal effect of blood serum was comprised of two components, one which had the ability to retain function following heating and the second which lost such effect after heat treatment (Walport, 2001, Morgan, 1990). The actual term 'complement' was then coined by Paul Ehrlich during the late 1890's as part of his larger proposals for the immune system, where he described the heat labile system to complement or aid the killing of bacteria by heat stable antibodies (Ehrlich, 1899).

The complement system itself, is comprised of over thirty proteins, including soluble serum proteins and some which are associated with the membrane (Tegla et al., 2011). Activation of complement triggers a sequential cascade of enzymatic reactions which in turn provides the platform to permit the formation of the anaphylatoxins c3a and c5a (Hugli, 1986) and is summarised in figure 1.23. These molecules have the ability to stimulate a range of physiological responses including opsonisation (Schifferli et al., 1986), chemotaxis of both macrophages and neutrophils (Arumugam et al., 2006, de Vries et al., 2003, Chenoweth & Hugli, 1978) modulation of the inflammatory response (Hugli, 1986), and finally cell apoptosis (Hugli, 1986).

Activation of complement has been described to arise through three different pathways, so named the classical, the alternative and the lectin pathway (Sarma & Ward, 2011). All three involve the sequential cleavage of various inactive zymogens until the formation of C3 where the three pathways converge as the activation products C3a, C3b and C5a are produced leading to the assembly of the membrane attack complex (Sarma & Ward, 2011).

Initiation of the alternative pathway is triggered by the presence of carbohydrates, proteins and lipids on non-self surfaces (Lambris et al., 2008). Low levels of C3b are continuously produced by C3 cleavage and available to bind to possible invading pathogens (Sarma & Ward, 2011). Such binding initiates Factor B recruitment to C3b followed by Factor C which cleaves Factor B producing the C3 convertase C3bBb (Sarma & Ward, 2011). The C3bBb molecule is stabilised by the presence of properdin within plasma which is released by activated neutrophils

(Kemper et al., 2010). Properdin or Factor P protein is a positive regulator of complement activation and acts to stabilise the C3 convertase binding to C3b thereby preventing cleavage by Factors H and I (Kemper et al., 2010). Recent data suggests that properdin has the ability to directly bind to both apoptotic and necrotic cells leading to complement activation (Kemper et al., 2010).

Initiation of complement via the classical pathway occurs when immune complexes are formed by the binding of IgG or IgM to non-self antigens including pathogens (Sarma & Ward, 2011). Following the formation of the C1 complex composed of C1q, C1r and C1 molecules, the C1 complex binds to the Fc region of the IgG or IgM immune complex (Sarma & Ward, 2011). This binding activates C1s which in turn cleaves C4 and C2 to form the CP C3 convertase (C4bC2a) (Sarma & Ward, 2011).

In contrast, the lectin pathway can be initiated in response to invading yeast, bacteria, parasites and viruses and involves the binding of the mannose binding lectin (MBL) or the molecule Ficolin to carbohydrate motifs on the surface of the invading pathogen (Sarma & Ward, 2011). MBL and Ficolin are in constant circulation within the host serum normally in complex with MBL-associated proteins known as MASPs (Sarma & Ward, 2011). One of particular importance is MASP2 which once activated following the binding to a pathogen becomes activated and stimulates the cleavage of C4 to form C4a and C4b (Wallis, 2007). The release and attachment of C4b to the pathogen surface induces the binding of C2 which is then available to be cleaved by MASP2 to form C2b and C2a (Sarma & Ward, 2011). The cleavage of these pair of molecules allows the formation of LP C3 convertase (C4BC2a) where the three independent pathways converge (Sarma & Ward, 2011).

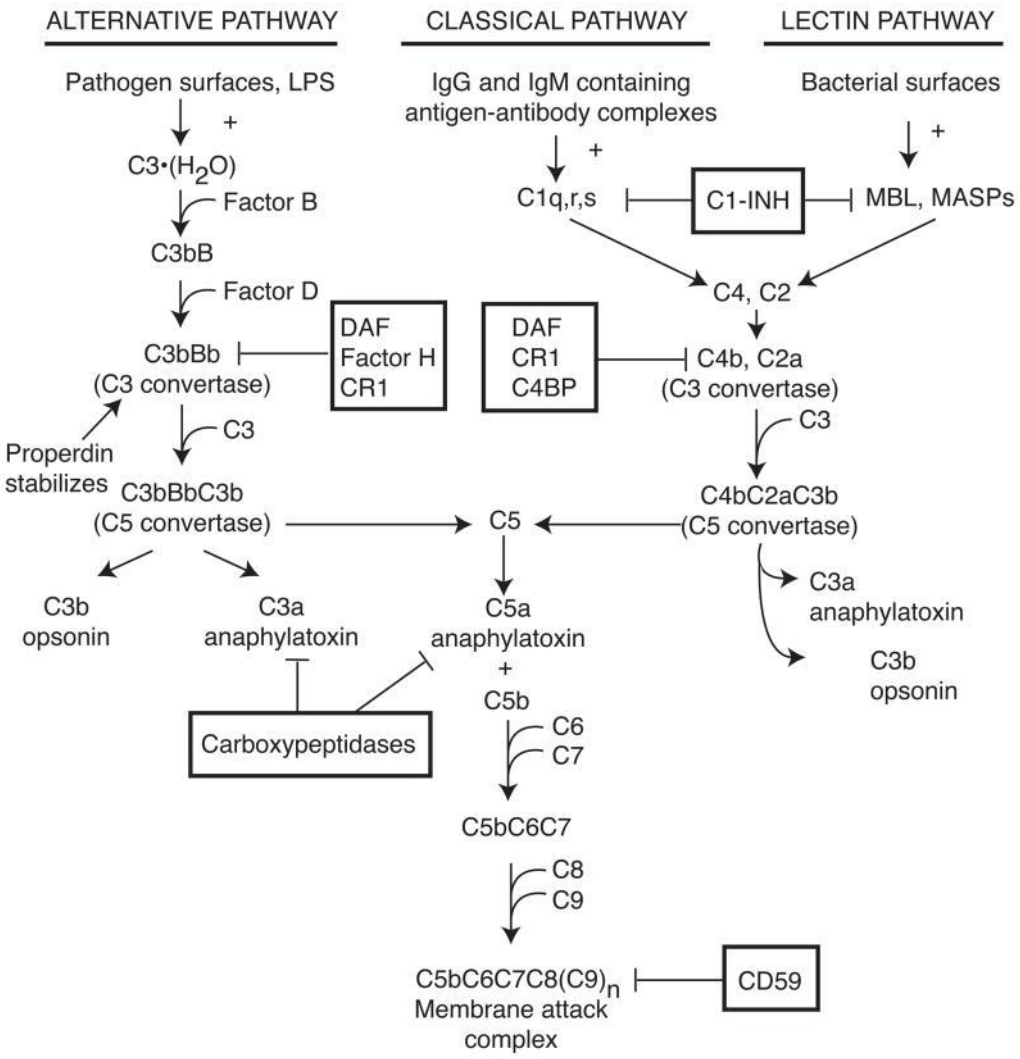


Figure 1.23. Complement activation pathways. The three different pathways of complement activation including the alternative, the classical and the lectin pathway. Molecules highlighted by boxes are factors known to inhibit or modulate the pathway. Adapted from Sarma and Ward, 2011.

Many pathogens have developed elegant and effective ways to prevent destruction by the complement response and are known to target every aspect of the system from the early stages of complement activation via any pathway to the inhibition of end results such as opsonisation, phagocytosis and cell lysis (Lambris et al., 2008). Bacterial pathogens are particularly adept to such evasion which is achieved by the proteolytic removal of complement proteins or by simple complex interactions with complement mediators such as human factor H and factor H like protein 1 (Lambris et al., 2008).

The *Borrelia* genospecies can be considered to vary with regards to their susceptibility to complement (Alitalo et al., 2001, Breitner-Ruddock et al., 1997, Pausa et al., 2003, Herzberger

et al., 2007). The group can also be further classified by their sensitivity or resistance to serum which is measured by determining the borreliacidal effect of the serum by means of the formation of the membrane attack complex upon the surface of the spirochaete as shown in figure 1.24. *B. garinii* serotypes 5 and 6 are classed as serum sensitive, *B. burgdorferi* as moderately serum resistant and *B. afzelii, B. spielmanii* and *B.bavariensis* (*B. garinii* serotype 4) as serum resistant (Alitalo et al., 2001, Breitner-Ruddock et al., 1997, Pausa et al., 2003 and Herzberger et al., 2007).

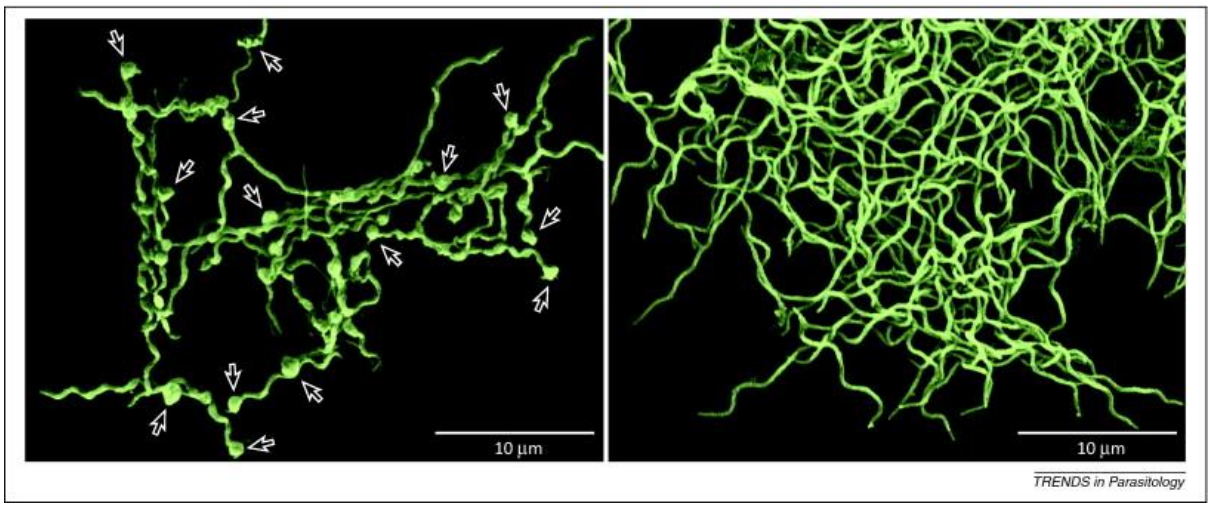


Figure 1.24. Complement destruction of *Borrelia* spirochaetes. Assessment of serum sensitivity of *Borrelia* genospecies by scanning electron microscopy following incubation with 12.5% human serum for 1.5 hours. **Left** - serum-sensitive *B. garinii* A87S with arrows highlighting bleb formation caused by complement activation and the formation of the membrane attack complex (C5b-9) leading to membrane disruption and cell death. **Right** – Serum-resistant *B. burgdorferi* B31 with no noticeable blebs highlighting the failure of complement activation. Micrographs taken from de Taeye et al., 2013.

In order for *Borrelia* to infect and establish an infection within the host the spirochaete must circumvent the host complement response both during and after transmission (de Taeye et al., 2013). During the early stages of transmission the tick salivary protein TSLPI (Tick salivary lectin pathway inhibitor) plays a key role in preventing complement activation. The tick protein inhibits the binding of mannose-binding lectin to its ligand and blocks complement activation via the lectin pathway (de Taeye et al., 2013). The addition of this tick protein to a serum incubation of serum-sensitive *Borrelia* strains also reduced complement mediated killing emphasising the importance of this tick protein during the early stages of transmission (de Taeye et al., 2013).

Serum resistant *Borrelia* genospecies have evolved complex mechanisms of avoiding complement mediated destruction (de Taeye et al., 2013), however these mainly involve the procurement of host complement regulators on to their surfaces (Lambris et al., 2008). The main regulator for the activation of the alternative pathway is the plasma component factor H which is a large glycoprotein (150kDa) consisting of twenty complement control protein modules (CCPs) each spanning approximately sixty residues (Schmidt et al., 2008). This regulation is essential for the differentiation between self and non-self cells and often recruited by bacterial pathogens. This targeted recognition of microbial surfaces as opposed to activation on host surfaces is crucial and tight regulation is required to prevent uncontrolled complement activation. The ability of factor H to recognise host and non-host surfaces can be demonstrated by mutations in the carboxyl-terminus of the glycoprotein. Mutations within this region can cause unchecked alternative pathway activation upon the surfaces of host cells leading to often irreparable damage to erythrocytes, platelets and endothelial cells. This damage may result in serious systemic disease such as atypical haemolytic uremic syndrome (Jokiranta et al., 2007).

Much research has concentrated around the expression of complement regulator acquiring surface proteins (CRASPs) by *Borrelia* species with crystal structures deposited within the PDB for much of the protein family. CRASPs are lipoproteins and are abundantly expressed on the outer surface of *Borrelia* (Kraiczy et al., 2001, Hallström et al., 2010, de Taeye et al., 2013). They act to prevent the complement cascade by binding to at least one member of the host factor H family of proteins (Kraiczy et al., 2001, Hallström et al., 2010, de Taeye et al., 2013). This acquisition of host factor H and factor H like proteins reduces the levels of complement activation by the alternative pathway, which results in the prevention of the formation of the membrane attack complex upon the spirochaetes surface (Kraiczy et al., 2001 and Bykowski et al., 2008). The factor H family consists of complement factor H itself, factor H-like protein (FHL-1), a splice variant of factor H and five complement factor H-related proteins (CFHR 1-5) with roles which are poorly understood but may act as cofactors for factor H. In turn there are five known CRASPs (1-5) also known as Erp's which are summarised in table 1.4 (Kraiczy et al., 2001, Hallström et al., 2010, de Taeye et al., 2013, Pulzova et al., 2014). The five CRASPs have all been identified in serum resistant *Borrelia* strains across *B. afzelii* (BaCRASPs) and *B. burgdorferi* (BbCRASPs) with factor H binding partners summarised in table 1.9 (de Taeye et al., 2013).

	Factor H	FHL-1	CFHR-1	CFHR-2	CFHR-3	CFHR-4	CFHR-5
BbCRASP-1	Y	Y	-	-	-	-	-
BbCRASP-2	Y	Y	-	-	-	-	-
BbCRASP-3	Y	-	Y	Y	-	-	Y
BbCRASP-4	Y	-	Y	Y	-	-	Y
BbCRASP-5	Y	-	Y	Y	-	-	Y

Table 1.9. *B. burgdorferi* CRASPs and factor H binding. A summary of the known interactions between *B. burgdorferi* CRASPs and Factor H or factor H like proteins. CRHR1-5 are known as complement factor H related proteins. Y indicated an interaction between the *Borrelia* protein and the complement regulator.

1.6.4 Surface antigenic variation

Antigenic variation remains a common mechanism of immune evasion for many bacteria with the most extensively studied organisms including, *Neisseria meningitides* and *N. gonorrhoeae* (Deitsch et al., 2009). Recombinational shuffling of *B. burgdorferi's* genetic material (further discussed in section 1.3.4) allows surface epitopes to be continually refreshed (Zhang & Norris, 1998). This progressive change to the pathogens outer surface frustrates the host's humoral immune response and provides the spirochaete with an ever changing disguise.

The outer surface lipoprotein VlsE (Variable major protein/Vmp-like sequence) from *B. burgdorferi* is known to undergo rapid antigenic variation and is believed to play a crucial role in evading the host's immune response (Zhang & Norris, 1998). Structurally the 35-kDa protein contains six variable regions which form loop structures and occupy the membrane surface whilst covering the invariant regions of the protein (Eicken et al., 2002). Although the function of VlsE is poorly understood it is known that the protein is significant during mammalian infection and the loss of expression by the removal of the coding region from it's plasmid Ip28-1, results in reduced infectivity (Barbour, 1984 and Barbour et al., 1986). Further evidence supporting the importance of VlsE is that mutant *Borrelia* that lack the plasmid Ip28-1 are incapable of long term mammalian infection (Labandeira et al., 2003 Labandeira and Skare, 2001, Purser and Norris, 2000).

Antigenic variation is achieved by use of the *vls* locus which is comprised of the *vlsE* expression site followed by 11 to 15 silent *vls* cassettes (Zhang & Norris, 1998). By recombining segments of the silent cassettes with the *vlsE* region the resulting *vlsE* gene and the eventual outer surface protein is highly variable at specific surface exposed regions

protecting the invariant core (Liang et al., 2000, Zhang & Norris, 1998, Eicken et al., 2002). An important note is that this variation is only seen during mammalian infection and not during the colonisation of ticks or *in vitro* cultivation. In terms of speed, antigenic variation of *vlsE* seen in mice shows detectable sequence changes within 4 days and by 28 days every isolate can be considered unique with between 9-13 recombination events (Zhang and Norris, 1998). Similar systems have been identified in other pathogens associated with persistent infection and relapsing fever and include the protozoa *Plasmodium* spp, *Trypanosoma* spp (Barbour and Restrepo, 2000) and relapsing fever spirochaetes (Schwan and Hinnebusch, 1998).

Expression studies have also demonstrated that *B. burgdorferi* keeps expression of *VlsE* under tight control and is seen to produce little when colonising the midgut of the tick and upregulates expression during mammalian infection (Indest et al., 2001) indicating that expression is most probably controlled in line with both exogenous and endogenous signals associated with the natural infectious cycle.

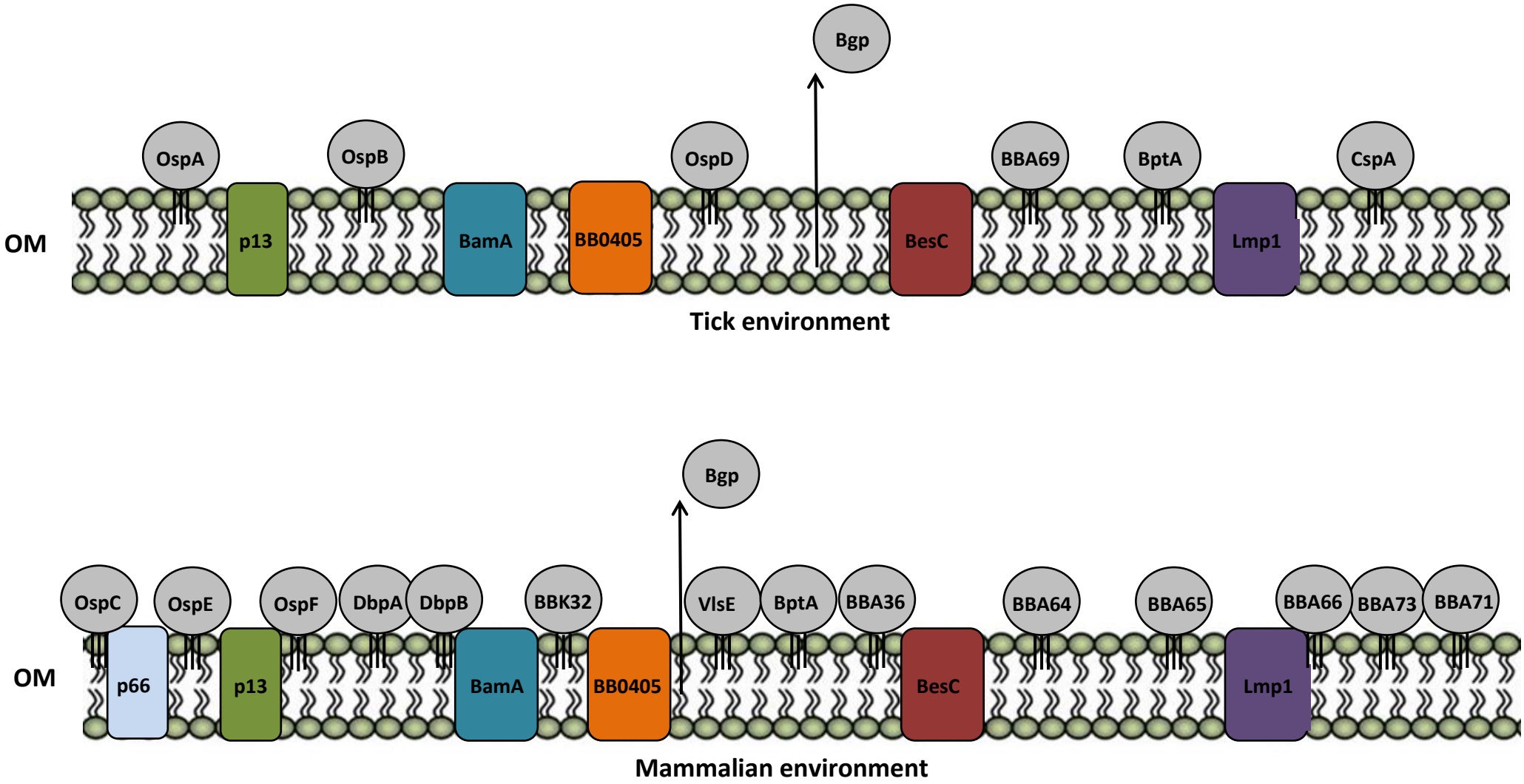


Figure 1.25. Outer membrane protein regulation during tick and mammalian infection. The profile of the *Borrelia* outer membrane is seen to drastically shift between hosts. Some proteins are specifically associated with expression during tick infection whereas others are notably involved during mammalian infection. Although OspA is in higher abundance during tick infection the protein has also been implicated in roles during mammalian infection and can be considered to be present during both environments.

1.6.5 Oxidative stress and *Borrelia* superoxide dismutase

B.burgdorferi's ever changing environment often causes exposure to fluctuations in O_2 and challenges by the host's immune system contributes further to the oxidative stress placed upon the spirochete (Wang et al., 2013). In order to survive the hostile environment *B.burgdorferi* can exploit a tick salivary gland protein, Salp25D when migrating from arthropods to mammals and although the spirochete lacks peroxidase or catalase enzymes its genome encodes a single superoxide dismutase gene, *sodA (bb0153)* which is essential for virulence (Esteve-Gassent et al., 2008).

Superoxide dismutases (SODs) are a group of crucial enzymes which catalyse the disproportionation of superoxide to hydrogen peroxide and oxygen. Enzymatic activity requires a metal ion co-factor which can be used to distinguish between the different SOD isoforms which include Mn-SOD, Fe-SOD, Cu/Zn-SOD, Ni-SOD and cambialistic SODs which can effectively function with either manganese or iron as its co-factor (McCord & Fridovich, 1969, Yost & Fridovich, 1973, Keele & McCord, 1970, Youn et al., 1996, Martin et al., 1986). The different isoforms can be experimentally identified by their sensitivity to cyanide (Cu/Zn-SOD), sensitivity to hydrogen peroxide (Fe-SOD and Cu/Zn-SOD) or resistant to both cyanide and hydrogen peroxide treatments (Mn-SOD) (Asada et al., 1975, Benov & Fridovich, 1994). The family of Mn and Fe-SODs are particularly well conserved from bacteria to humans and Mn and Fe versions share significant homology and can bind either co-factor with similar affinities and binding geometries (Mizuno et al., 2004, Iranzo, 2011, Kang et al., 2011, Yamakura et al., 2007). Although the binding affinities of both Fe and Mn-SODs for their co-factors are similar the binding of the incorrect metal ion renders the enzyme inactive possibly due to the disruption of redox potential over any physical changes to conformation (Vance and Miller, 1998, Jackson and Brunold, 2004).

SodA from *B. burgdorferi* was originally characterized as a Fe-SOD based upon its resistance to cyanide yet enzymatic sensitivity to H_2O_2 (Whitehouse et al., 1997). Three years later a study demonstrated that this classification for *Borrelia* was unlikely, their reasoning included that; *Borrelia* have an extremely low intracellular iron concentration with evidence suggesting this could be as low as 10 atoms per cell, there is no growth requirement for iron and there is a distinct lack of cellular machinery for the uptake of iron (Posey and Gherardini., 2000). Further studies failed to confirm Whitehouse's findings that *B.burgdorferi's* SOD was enzymatically sensitive to H_2O_2 but resistant to cyanide. However, instead suggested that the *B.burgdorferi* SOD was resistant to both treatments and was indeed a Mn-SOD (Troxell et al., 2012). The

group also suggested that the enzyme was unlikely to be cambialistic due to the low intracellular iron content and the lack of iron uptake proteins encoded within the genome.

Within metalloprotein biology or evolutionary studies *B. burgdorferi* can be classed as a particularly interesting organism as it has evolved with no requirement for iron (Posey and Gherardini., 2000, Fraser et al., 1997). As previously described not only is the intracellular iron concentration extremely low the spirochaete can rapidly acquire remarkably high levels of manganese (Aguirre et al., 2013). Metalloproteomic studies have demonstrated that particularly high levels of manganese are crucial for SodA activation within *B. burgdorferi*. The study identified that the majority of the cells SodA enzyme associates with manganese with a smaller subset being held as inactive apoprotein (Aguirre et al., 2013). The study also demonstrated that although *B. burgdorferi* has evolved to a manganese rich and iron poor environment the opposite is seen for other microorganisms such as *Saccharomyces cerevisiae* and *E. coli* where the intracellular concentration of iron is high and the manganese co-factor concentration is low yet manganese is still captured by these SODs. In order to determine whether the *B. burgdorferi* SodA behaves in a similar manner the enzyme was expressed in the iron-rich mitochondria of S. cerevisiae and was found to be inactive suggesting the *B. burgdorferi* SodA is only active in the presence of a high manganese concentration.

Recent modelling studies of *B. burgdorferi* SodA have highlighted strong homology to other known Mn-SODs and computer aided modelling of the proposed active site using the modelling software, Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000) has also highlighted some possibly unusual features within the hydrogen bonding near the enzymes active site which are not shared with other known bacterial SODs such as E. coli SodA (Lah et al., 1995). These differences are likely to have occurred due to *B. burgdorferi* adaption to a manganese rich and iron poor environment which is likely to give the pathogen a competitive edge when the infected host begins to starve the invader of iron (Posey and Gherardini, 2000, Correnti and Strong, 2012 and Johnson and Wessling-Resnick, 2011). The single *B. burdorferi* SOD gene has also been demonstrated to be essential for virulence highlighting the importance of characterisation of this normally well conserved enzyme.

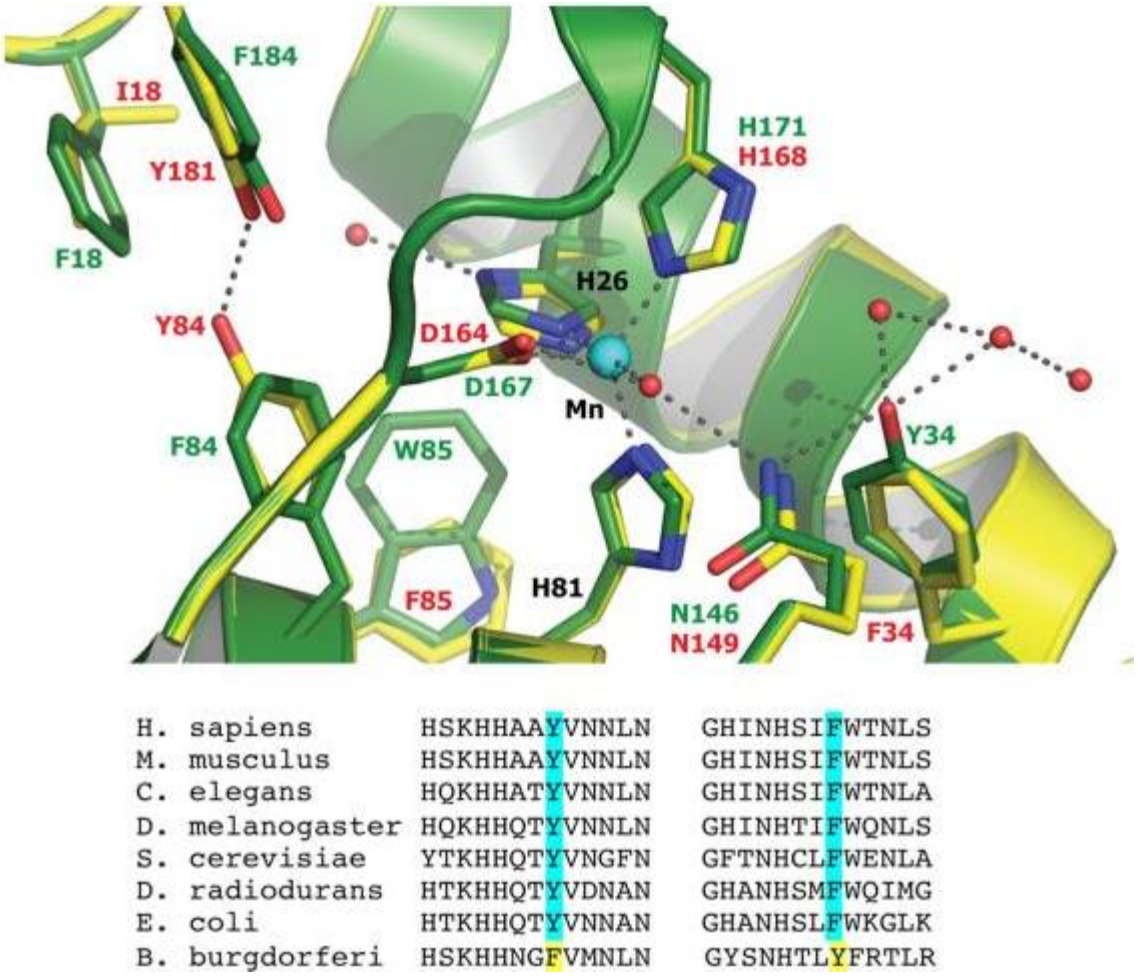


Figure 1.26. Predicted active site of *B. burgdorferi* SodA. Homology models for *B. burgdorferi* SodA were generated using the above known Mn-SOD templates and generated using Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000). The figure shows the predated active site taken from Aguirre et al., 2013. The yellow represents residues from *B. burgdorferi* SodA and the green denotes residues from *E. coli* SodA.

1.7 Research Aims

1.7.1 *Borrelia* OmpA-like proteins

The *B. burgdorferi s.l.* genome is known to encode a large number of lipoproteins with upto 8% of genes implicated. Much research has revolved around the identification and characterisation of lipoproteins with much progress made. However, many of these are often encoded on *Borrelia's* vast number of plasmids which are often variable between genospecies (Fraser et al., 1997). This variability of plasmids between *Borrelia* strains has hindered the development of new vaccines and diagnostic methods. Freeze fracture electron microscopy has highlighted that *Borrelia s.l.* has an unusually low abundance of integral OMPs, possibly upto 10-fold lower than what can be detected from an *E. coli* outer membrane (Lugtenberg and van Alphen, 1983, Radolf et al., 1994). This lower abundance alongside the limited antigenicity of known OMPs has possibly stalled the identification and characterisation of these integral proteins and in stark contrast to lipoproteins many of these hypothetical integral membrane proteins are encoded on the stable chromosome (Fraser et al., 1997).

To date approximately ten OMPs have been identified in *B. burgdorferi* (Brooks et., 2005, Lenhart & Atkins., 2009, Antonara et al., 2007, Bunikis et al., 2008, Coburn & Cugini et al., 2003, Cugini et al., 2003, Noppa et al., 2001, Skare et al., 1997, Wood et al., 2013 and Russell & Johnson., 2013) and it is likely that these ten proteins represent only a small subset of these *Borrelia's* exposed integral proteins with many still not identified. Several studies have highlighted the importance of the identification and characterisation of these novel proteins (Tan et al., 2008, Wang et al., 2009, Jolley et al., 2001, Okamura et al., 2012) with the ultimate goal that some may offer suitable vaccine targets (Dunn et al., 2015) or aid current diagnostic tests.

As the eight stranded membrane spanning beta barrel, OmpA from *E. coli* is highly conserved within Gram-negative bacteria it seemed appropriate to use this key protein as the basis of the initial search for novel OMPs in *Borrelia*. The OmpA protein has multiple cellular functions with involvement in, adhesion, structural integrity, immune evasion, biofilm formation and has also been associated with invasion of brain microvascular endothelial cells highlighting the importance of identifying homologs within other pathogenic bacteria.

Previous work by Dr Adam Dyer (Identification and Structural Characterisation of Novel Outer Membrane Proteins in *B. burgdorferi s.l.* – Thesis 2013, University of Huddersfield) identified

BAPKO_0422 and a number of paralogous genes across *B. burgdorferi*, *B. afzelii* and *B. garinii* which were predicted to encode proteins believed to form 8-stranded outer-membrane β-barrels. This identification was based upon a novel approach using hidden Markov models and fold prediction as traditional protein BLAST (Altschul et al., 1990) searching failed to identify any characterised proteins with any significant degree of sequence similarity. The identification of integral membrane proteins in this manner is often hindered due to the substantial sequence divergence of these proteins particularly within the sequences associated with the external loop regions which are highly variable. The research undertaken by Dyer, 2013 employed a unique approach to identify possible OmpA-like proteins based upon the creation of HMM profiles of known OmpA-like proteins and using this profile to search the *Borrelia* genome followed by further investigation using fold and signal sequence prediction. For full methodology and results please refer to the original text (Dyer, 2013, Dyer et al., Appendix 4). Following this bioinformatics study a number of proteins were identified as possible OmpA-like proteins and are summarised in table 1.10.

Ordered locus name	**GenBank (UniProtKB) Accession codes**	**No. of amino acids**	**FFAS03 highest scoring template**	**Signal sequence prediction (SignalP 4.1)**
BAPKO_0026	ABH01291.1 (Q0SPD7)	211	OmpW (2X27)	N
BB_0027	AAC66429.1 (O51058)	212	OmpW (2X27)	Y
BG0027	AAU06886.1 (Q662Y5)	212	OmpW (2X27)	N
BAPKO_0422	ABH01676.1 (Q0SNA2)	201	OmpA (2KOL)	Y
BB_0405	AAC66795.1 (O51366)	203	OmpW (2X27)	Y
BG0407	AAU07257.1 (Q661L4)	203	OmpW (2X27)	Y
BAPKO_0423	ABH01677.1 (Q0SNA1)	203	OmpA (2KOL)	Y
BB_0406	AAC66794.1 (O51367)	203	OmpA (2KOL)	Y
BG0408	AAU07258.1 (Q661L3)	203	OmpA (2KOL)	Y
BAPKO_0591	ABH01831.1 (G0IQB2)	181	NspA (1P4T)	Y
BB_0562	AAC66924.1 (O51510)	180	NspA (1P4T)	N
BG0572	AAU07409.1 (Q660W2)	179	NspA (1P4T)	Y

Table 1.10. Possible OmpA-like proteins identified by HMM/FFAS03/BLAST. Proteins were identified by searches from *B. afzelii, B. burgdorferi* and *B. garinii* translated open reading frames using hidden Markov models and signal sequence analysis (Dyer, 2013).

Following the initial identification a number of the possible 8-stranded beta barrel proteins were found within past literature and included BG0407 which had been noted as a human factor H binding protein (Bhide et al., 2009) and BB0405 which had previously been identified as a monomeric protein in outer membrane vesicles (Yang et al., 2011).

The overall aims of the research was to clone a selection of these hypothetical proteins across all three *Borrelia* strains and produce recombinant protein in order to probe their structural and functional relationships and determine whether these proteins were indeed part of the OmpA-like domain family.

Summary of the research aims - *Borrelia* OmpA-like proteins:

- Clone and produce recombinant OMPs from across *B. burgdorferi*, *B. afzelii* and *B. garinii* based upon the initial bioinformatic identification of possible OmpA-like proteins.
- Determine possible binding motifs within these proteins using bioinformatic techniques.
- Identify possible binding ligands *in vitro*.
- Explore possible conformations by means of homology modelling.
- Characterise the unknown proteins using phase partitioning, circular dichroism, small angle X-ray scattering and X-ray crystallography.
- Conclude whether these proteins are likely to belong to the OmpA-like domain family.

1.7.2 Superoxide dismutase A

Alongside the investigation of possible *Borrelia* OMPs it was also planned to investigate another crucial protein, *B. burgdorferi* superoxide dismutase A. The small enzyme was selected for further investigation as it lacked full characterisation with regards to its enzymatic activity and its structure was yet to be solved although the protein is considered highly conserved among other bacteria.

The complex lifecycle of *B. burgdorferi* often causes exposure of the pathogen to fluctuations in oxygen levels and the potential for the spirochaete to encounter reactive oxygen species (ROS) (Wang et al., 2013). Previous studies have identified that *B. burgdorferi's* superoxide dismutase A (SodA) (23.5kDa protein) is the key metalloenzyme responsible for degradation of ROS. Early work suggested that the enzyme required the presence of iron to function (Whitehouese et al., 1997). However, it is now postulated that the required co-factor is manganese, as *B. burgdorferi's* intracellular concentration of iron is minimal, at less than 10 atoms per cell (Posey et al., 2000), there is a distinct lack of iron containing proteins and the spirochaete does not have the capacity to transport iron nor require it for growth (Radolf et al., 2012).

Much work has been published supporting that this enzyme is indeed a manganese SOD (Posey and Gherardini., 2000, Troxell et al., 2012, Radolf et al., 2012). However, no studies to date have identified whether this enzyme is cambialistic, whereby it could use both manganese and iron as its co-factor.

Summary of the research aims - *B. burgdorferi* superoxide dismutase A:

- Produce recombinant superoxide dismutase A
- Determine the metal co-factor required for enzymatic activity *in vitro* and determine whether the protein exhibits cambialistic behaviour.
- Produce homology models.
- Characterise the protein using circular dichroism, small angle X-ray scattering and X-ray crystallography.

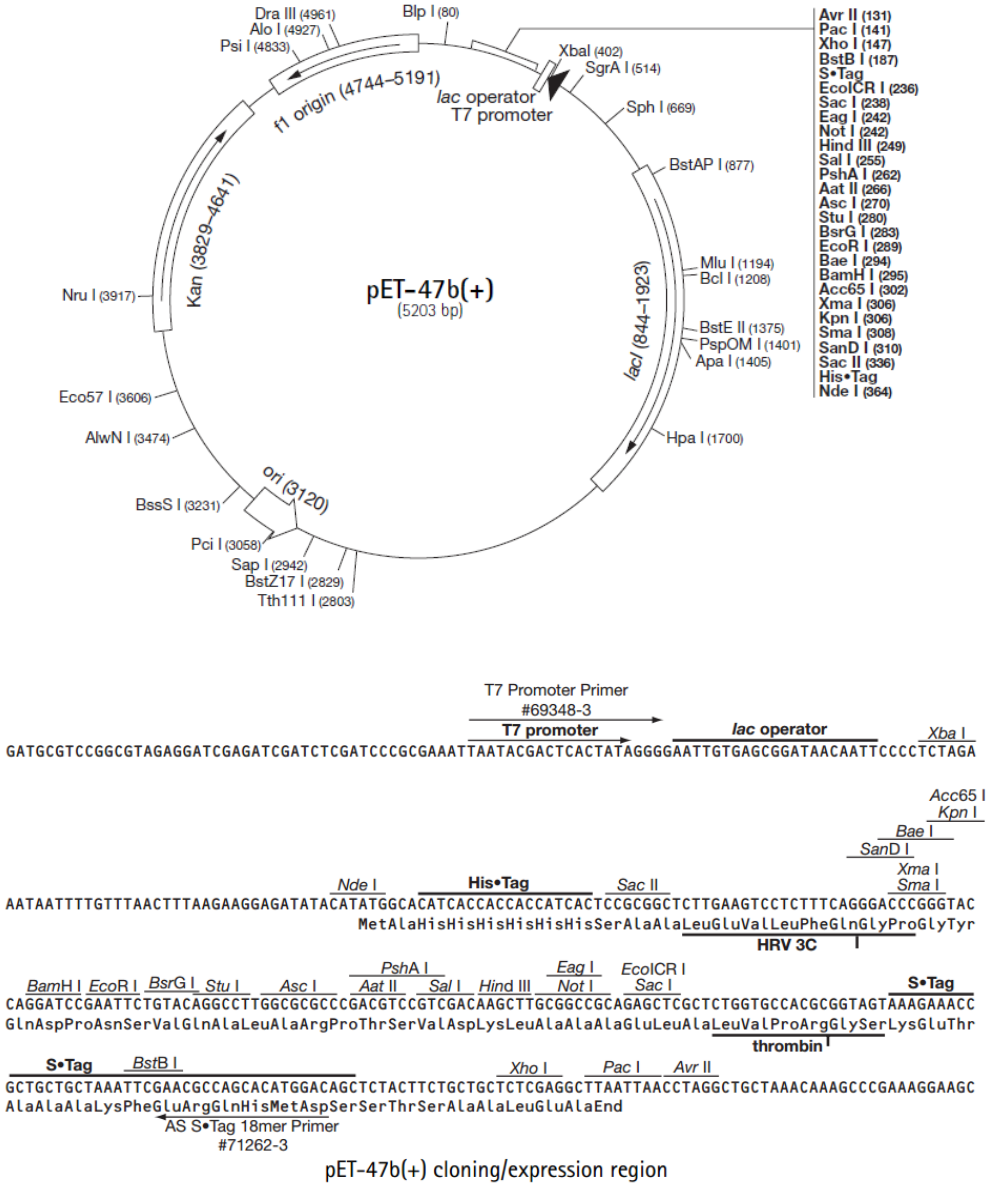
Chapter 2: Materials and Methods

2.1 Materials

2.1.1 DNA and Oligonucleotides

The BAPKO_0422-pET47 expression construct was kindly supplied by Dr Adam Dyer (The University of Huddersfield). *Borrelia* genomic DNA used for the development of further constructs was supplied by Dr Gabi Margos (The University of Bath), from *Borrelia* DNA extracted from infected *I.rinicus* and also from Dr Volker Fingerle (German National Reference Centre for *Borrelia*) who provided purified genomic DNA from *B. burgdorferi s.s B31, B. garinii PBi* and *B. afzelii PKo* strains. The *B. burgdorferi* SodA construct was commercially codon optimized for expression within *E. coli* and synthetically produced by MWG Operon, Eberberg, Germany.

The pET-47b(+) DNA was purchased from Novagen (Merck Serono Ltd, Middlesex, UK) and was used for targeted gene expression with the plasmid map shown in figure 2.1. The plasmid contains a T7 promoter and terminator along with the coding sequence for an N-terminal histidine tag followed by the recognition sequence for the highly specific (HRV) 3C protease. The presence of the 6-his-tag provided an ideal platform for immobilised metal affinity chromatography purification and the plasmid also conferred kanamycin resistance for selection.



pET-47b(+) cloning/expression region

Figure 2.1 - pET-47b(+) Plasmid map and cloning region – Overview of the vector map and the nucleotide sequence of the cloning region with restriction sites. (Taken from Novagen®)

Primers used throughout this project were designed and analysed using OligoExplorer 1.2 and OligoAnalyzer 1.2 (Genelink). Primers were synthesised by Eurofins MWG Operon (Germany) and provided as lyophilised samples which were diluted to 100pmol/µl following manufacturer instructions using sterile nuclease free water (full oligonucleotide data sheets can be viewed within the appendix). Primers used to target genes *bb_0562, bb_0406* and *bg0408* are summarised in table 2.1. Genes and proteins targeted are referred to using their Uniprot identifiers.

Initial amplification primers were designed to target the gene of interest plus an additional ~10 base pairs both up and downstream of the region. Restriction primers were then used to reduce this region back to its native sequence and to also add the required restriction site. Prior to selection of a restriction site the sequences were submitted to the web server NEBcutter v2.0 in order to determine whether any restriction sites were naturally present within the nucleotide sequence, output data is shown in figure 2.2. Following the identification of any unsuitable restriction sites within the three genes, BamHI and NotI were selected as suitable restriction sites to add to the generated PCR products as shown in table 2.2. Neither of these two restriction sites were naturally present within the three genes nor present at any point with the vector apart from within the multiple cloning site.

Target	Primer Name	Sequence (5'-3')	Tm (ºC)	Length
BB_0562	BB_0562_F	TAT AATAATAAGACAAAATTTGGGAG	53.8	26
	BB_0562_R	GAATTTTTTCATAGCGCTTGGC	56.5	22
BB_0406	BB_0406_F	CTTTATAAAAATATATATATAGTGGAG	52.8	27
	BB_0406_R	TATGCGAGCTTAATTTTAATTACC	54.2	24
BG_0408	BG0408_F	TTAAGTAAGCTTAATTTTA ATCACC	53.1	27
	BG0408_R	TATAATAATAAGACAAAAT TTGGGAG	53.8	25

Table 2.1 - Primers for initial amplification – Primer sequences, lengths and melting temperatures for oligonucleotides used to generate initial PCR products for expression construct development. *Borrelia* strains are denoted by **BB**, *B. burgdorferi* B31, **BG**, *B. garinii* PBi, **BAPKO**, *B. afzelii* PKo followed by the annotated gene number as recorded on NCBI (https://www.ncbi.nlm.nih.gov/) and the primer type of either F, forward or R, reverse.

Target	Primer Name	Sequence (5'-3')	Tm (ºC)	Length
BB_0562	BB_0562_F_BamHI	AATTATGˇGATCCGAAAGATTCATATTTA AATAGAGGA	62.8	37
	BB_0562_R_NotI	ATAAGCˇGGCCGCGAATTTTTTCATAGCG CT	66.8	30
BB_0406	BB_0406_F_BamHI	AATTATGˇGATCCGTCTGACAATTATATG GTCAG	64.5	33
	BB_0406_R_NotI	ATAAGCˇGGCCGCTATGCGAGCTTAATTT TAATTACC	68.3	36
BG_0408	BG0408_F_BamHI	AATTATGˇGATCCGTCTGACAATTATAT GGTCAG	64.5	33
	BG0408_R_NotI	ATAAGCˇGGCCGCTTAAGTAAGCTTAAT TTTAATCACC	67.2	37

Table 2.2 - Restriction site primers – Primers were designed to add appropriate restriction sites and to keep the amplified fragment in a correct reading frame for expression. Between 4 to 6 bases were added to the ends of the oligonucleotides to improve restriction digestion efficiency.

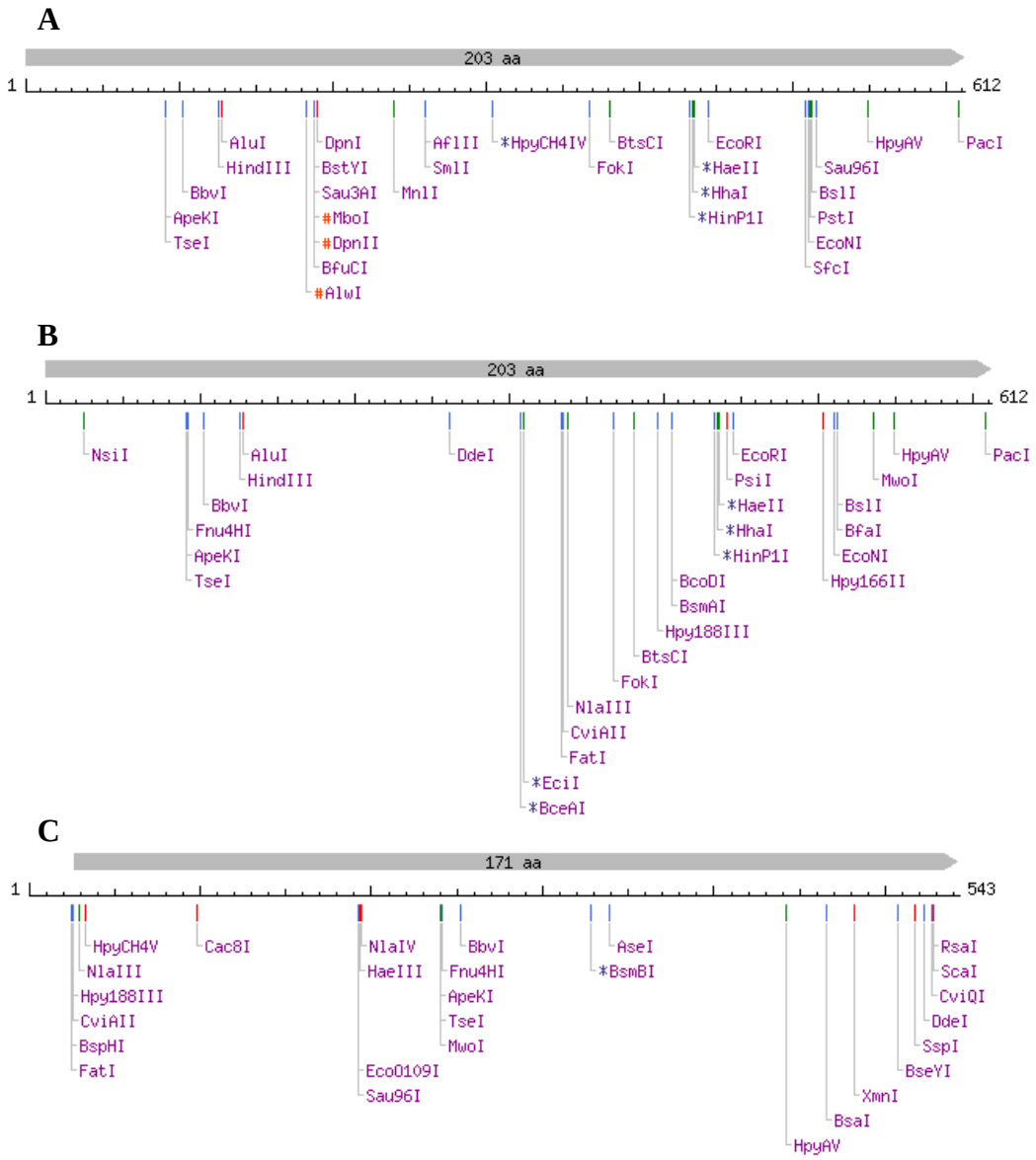


Figure 2.2 – Restriction site analysis of *Borrelia* OMP genes to be cloned. Full graphical output highlighting the naturally present restriction sites in each of the three *Borrelia* OMP genes, generated using NEBcutter v2.0. From top to bottom, A - BB_0562, B - BB_0406 and C - BG0408. Red cut sites represent blunt ended cutting, blue cut sites indicate a 5' extended cut, green a 3' extended cut and black bars cuts one strand only.

2.1.2 Reagents, buffers and enzymes

Required chemicals and reagents were purchased from Sigma-Aldrich (Gillingham, UK) and Fisher Scientific (Loughborough, UK), DNA purification kits from Qiagen and antibodies from Abcam.

Chemical/Reagent	Supplier	Abbreviation
Acrylamide (40% in solution)	Fisher	Bis-tris
Agar	Fisher	-
Ampicillin sodium salt	Fisher	Amp
Ammonium persulfate	Sigma	APS
Blue Dextran	Sigma	-
Bovine serum albumin	Sigma	BSA
Bradford reagent	Invitrogen	-
Chloramphenicol	Fisher	Chl
Complement human Factor H	Cambridge Bioscience	hFH
Dithiothreitol	Fisher	DTT
Ethylenediaminetetraacetic acid	Fisher	EDTA
Glycerol	Fisher	-
Guanidine Hydrochloride	Fisher	-
Imidazole	Fisher	-
Iron(II) chloride	Fisher	$FeCl_2$
Isopropyl β-D-1-thiogalactopyranoside	Fisher	IPTG
Kanamycin sulphate	Fisher	Kan
Lauryldimethylamine-oxide	Sigma	LDAO
Lysozyme	Sigma	-
Manganese(II) chloride	Fisher	$MnCl_2$
Nickel(II) sulphate	Sigma	-
Protease inhibitor cocktail	Sigma	-
Ribonuclease A	Sigma	-
Sodium azide	Fisher	NaN_3
Sodium dodecyl sulphate	Sigma	SDS
Sodium chloride	Fisher	NaCl
Sodium phosphate monobasic	Fisher	-
Sodium phosphate dibasic	Fisher	-
Sucrose	Sigma	-
Tetramethyethylenediamine	Sigma	TEMED
Tris (hydroxymethyl)aminomethane-base	Fisher	Tris-Base
Tris (hydroxymethyl)aminomethane-HCl	Fisher	Tris-HCl
Triton X-100	Sigma	-

Triton X-114	Sigma	-
Tryptone	Fisher	-
Tween-20	Sigma	-
Urea	Fisher	-
Yeast extract	Fisher	-
1-oleoyl-rac-glycerol	Sigma	Monoolein
3-[(-Chloamidopropyl)dimethylammonio]-1-propanesulphonate	Fisher	CHAPS

Table 2.3 - Chemicals and reagents – A list of chemicals and reagents used throughout the project and the supplier.

Enzyme	**Reaction buffer**	**Restriction site/ site of activity**	**Supplier**
BamHI FastDigest	FastDigest Buffer	G^GATCC CCTAG^G	Thermo
DreamTaq DNA Polymerase	DreamTaq Buffer	___	Thermo
EcoRI	EcoRI Buffer	G^AATTC CTTAA^G	NEB
NotI FastDigest	FastDigest Buffer	GC^GGCCGC CGCCGG^CG	Thermo
SalI	NEBuffer 3 + BSA	G^TCGAC CAGCT^G	NEB
Taq Polymerase	Standard Taq Buffer	___	NEB
T4 DNA Ligase	T4 DNA Ligase Buffer	___	Thermo
HRV(3C) protease	HRV(3C) reaction buffer	Leu-Glu-Val-Leu-Phe-Gln-↓-Gly-Pro	Thermo

Table 2.4 - Enzymes – Restriction and cloning enzymes used throughout the project. When commercial HRV(3C) wasn't used the protease was expressed and purified in house using a H-3Cpro construct containing the protease gene with an N-terminal his-tag and MBP C-terminal tag.

Buffer/Media	Composition
Autoinduction media	1% w/v tryptone, 0.5% w/v yeast extract, 1x NPS, 1x 5052
Coomassie blue stain	0.1% w/v Coomassie R250, 10% v/v glacial acetic acid, 40% v/v methanol
Coomassie blue destain	10% v/v glacial acetic acid, 20% v/v methanol
IEC denaturing binding buffer	8M urea, 20mM NaCl, 50mM Tris, pH8
IEC denaturing elution buffer	8M urea, 1M NaCl, 50mM Tris, pH8
IEC native binding buffer	20mM NaCl, 50mM Tris, pH8
IEC native elution buffer	1M NaCl, 50mM Tris, pH8
IEC refold buffer	1M urea, 20mM NaCl, 50mM Tris, pH8
Inclusion body wash buffer 1	0.3 NaCl, 50mM Tris, 1mM EDTA, 10mM DTT, 5% v/v Triton X-100, pH8
Inclusion body wash buffer 2	0.3 NaCl, 50mM Tris, 1% v/v Triton X-100, pH8
Inclusion body wash buffer 3	0.3 NaCl, 50mM Tris, pH8
IMAC denaturing binding buffer	8M urea, 0.3M NaCl, 50mM sodium phosphate, pH8
IMAC denaturing wash buffer	8M urea, 0.3M NaCl, 50mM tris, 20mM imidazole , pH8
IMAC denaturing elution buffer	8M urea, 0.3M NaCl, 50mM tris, 0.5M imidazole, pH8
IMAC native binding buffer	0.3M NaCl, 50mM tris, pH8
IMAC native wash buffer	0.3M NaCl, 50mM tris, 20mM imidazole , pH8
IMAC native elution buffer	0.3M NaCl, 50mM tris, 0.5M imidazole , pH8
IMAC refold buffer	1M urea, 0.3M NaCl, 50mM tris, pH8
Lysis buffer	0.3M NaCl, 50mM Tris, 0.4µg/ ml lysozyme, 0.1% v/v Triton X-100, pH8
MES running buffer	50mM MES, 50mM Tris base, 0.1% w/v SDS, 1mM EDTA, pH 7.3
NPS x20	1 M Na_2HPO_4, 1 M KH_2PO_4, 0.5 M $(NH_4)_2SO_4$
TAE x10	400mM Tris base, 200mM acetic acid, 10mM EDTA
TBE x10	1M Tris base, 1M Boric acid, 0.02M EDTA
Transfer buffer	1.45% w/v Tris, 7.2% w/v glycine, 20% v/v methanol
TBS x1	0.15M NaCl, 20mM Tris, pH 7.4
TBS-TWEEN	0.15M NaCl, 20mM Tris, 0.05% v/v TWEEN-20, pH 7.4
SAXS buffer 1	0.3M NaCl, 50mM Tris, 0.06% w/v LDAO, pH8
SAXS buffer 2	0.3M NaCl, 50mM Tris, 0.1% w/v LDAO, pH8
Size exclusion buffer	0.3M NaCl, 50mM Tris, 0.1% w/v LDAO, pH8
Super optimal broth (SOC)	2% w/v trptone, 0.5% w/v yeast extract, 10mM Nacl, 2.5mM KCl, 10mM $MgCl_2$, 20mM glucose.
Solubilisation buffer	8M urea, 0.3M NaCl, 50mM Tris, pH8
Solubilisation buffer – low salt	8M urea, 20mM NaCl, 50mM Tris, pH8
50x5052	25 % w/v α-lactose, 2.5 % w/v glucose, 10 % w/v glycerol

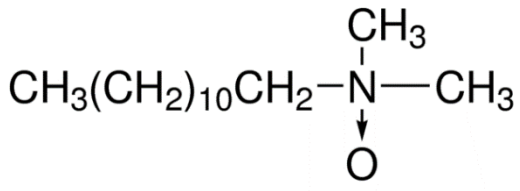
Table 2.5 -Buffers and media - All buffers used were prepared fresh, filtered (0.45µm) and de-gassed if required for chromatography.

2.1.3 Detergents

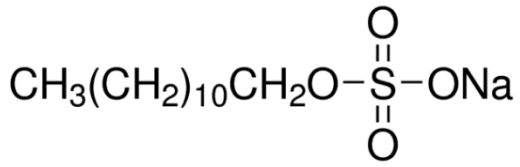
Detergents were purchased from Sigma Aldrich (Gillingham, UK). The detergents used throughout the project are summarised in table 2.6 and figure 2.3.

	Synonyms	CAS number	Molecular weight	Detergent type	Critical micelle concentration	Micelle size
LDAO	*N*,*N*-Dimethyldodecylamine *N*-oxide, LDAO, DDAO, Lauryldimethylamine *N*-oxide	1643-20-5	229.40	Zwitterionic	~1-2mM	17 kDa (Hermann, 1962)
SDS	Sodium dodecyl sulphate	151-21-3	288.38	Anionic	~8.2mM	18kDa
Triton X-100	4-(1,1,3,3-Tetramethylbutyl)phenyl-polyethylene glycol, t-Octylphenoxypolyethoxyethanol	9002-93-1	625	Non-ionic	~0.2-0.9mM	81kDa (Biaselle & Millar, 1975)
Tween -20	Polyethylene glycol sorbitan monolaurate, Polyoxyethylenesorbitan monolaurate	9005-64-5	1228	Non-ionic	~0.06mM	-----

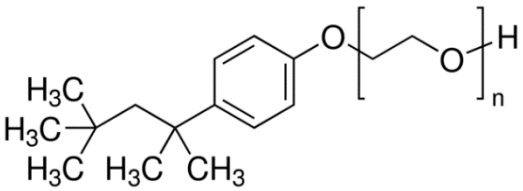
Table 2.6 – Detergents and their properties. A summary of the detergents used throughout the project. All detergents were purchased from Sigma Aldrich (Gillingham, UK).



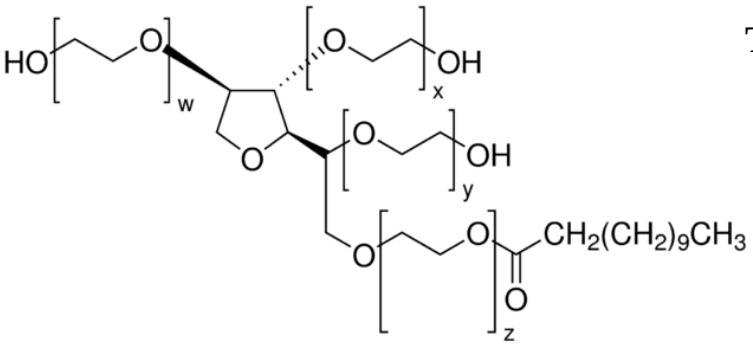
LDAO



SDS



Triton X-100



Tween-20

Figure 2.3 – Detergents and their chemical structures. The chemical organisation of the detergents used throughout the project.

2.1.4 Expression and maintenance *E .coli* strains

All *E. coli* strains were purchased from New England Biolabs or Merck Millipore as competent cells. NEB Turbo cells were used for molecular biology and for plasmid maintenance with BL21 and Rosetta cells used for protein overexpression. *E. coli* cells used throughout the project are summarised in table 2.7. Rosetta DE3 cells were utilized for expression of recombinant proteins that required rare codons not commonly found within the expression host *E. coli*. The additional tRNAs are supplied on a compatible chloramphenicol resistant plasmid and are driven by their native promoters.

Strain	Supplier	Usage	Genotype
E. coli Rosetta (DE3)	EMD Millipore	Expression	*F^- ompT $hsdS_B(R_B^- m_B^-)$ gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam^R)*
E. coli BL21 (DE3)	NEB	Expression	*F^- ompT gal dcm lon $hsdS_B(r_B^- m_B^-)$ λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*
E. coli T7	NEB	Expression	*fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10--Tet^S)2 [dcm] R(zgb-210::Tn10--Tet^S) endA1 Δ(mcrC-mrr)114::IS10*
E. coli Turbo	NEB	Maintenance	*F' $proA^+B^+$ $lacI^q$ ΔlacZM15 / fhuA2 Δ(lac-proAB) glnV galK16 galE15 R(zgb-210::Tn10)Tet^S endA1 thi-1 Δ(hsdS-mcrB)5*

Table 2.7 - *E. coli* strains used within the project with their genotypes – Both expression and maintenance strains were utilized within the project. Purchased cells were cultured for stock and subjected to 15% v/v glycerol (v/v) prior to flash freezing and storing at -80.

2.2 Methods

All commercial kits, enzymes and specialist reagents were used as recommended by the manufacturer.

2.2.1 Sequence analysis and BLAST searching

The amino acid sequences for all proposed OMPs were subjected to Signal P analysis (Petersen et al., 2011) in order to deduce the presence of signal peptides. Default parameters for D-cutoff values were used and the organism group selected was always Gram-negative. Protein BLAST (Altschul et al., 1990) searches were used to search for sequence similarity both within the *Borrelia* genus and throughout all other bacteria groups. All BLAST (Altschul et al., 1990) searching was performed using the NCBI (https://www.ncbi.nlm.nih.gov/) online tool using default parameters and BLOSUM62 (Henikoff & Henikoff, 1992) as the scoring matrix.

2.2.2 Molecular biology

2.2.2.1 PCR amplification of gene targets and addition of restriction sites

A PCR master mix was prepared for each set of amplification reactions. This contained all the required reagents with the exception of template DNA and primers. The volumes shown in table 2.8 were multiplied for each required reaction plus three additional, one for a negative control, the second a positive (a housekeeping gene) and one extra to account for pipetting error. The genomic DNA supplied by Dr Volker Fingerle (German National Reference Centre for *Borrelia*) was used at a constant of 20ng per reaction. All reagents were kept on ice prior to amplification in a Techne TC-312 thermocycler. The programmed cycle used was: 1 cycle of 2 minutes at 95^{o}C followed by 35 cycles of, 1 min at 94^{o}C, 30 to 60 seconds at a suitable annealing temperature and 1 min at 72^{o}C. The final extension step was at 72^{o}C for 5 minutes.

Reagent	Volume
$MgCl_2$	4 µl
X5 PCR buffer	10 µl
dNTPs	1 µl
Forward primer	1 µl
Reverse primer	1 µl
DNA template	2 µl
Taq Polymerase	0.25 µl
Sterile, nuclease free H_2O	30.75 µl
TOTAL VOLUME	50 µl

Table 2.8 A typical PCR reaction - A 50 µl PCR reaction volume was used for all stages of PCR, the total reaction volume was multiplied to cover all required reactions plus additional for controls and a spare.

Following agarose gel analysis and mini-prep of initial PCR products (Qiagen PCR purification kit), suitable restriction sites were added by PCR. 2µl of PCR product from the first amplification was added to a master mix as previously described, alongside 1µl of each restriction primer. The programmed cycle was designed in line with the primer melting temperatures and the number of cycles varied from 28 to 35 in order to maximise final yield.

2.2.2.2 Agarose gels

PCR products at various stages of the experimental procedure were analysed by agarose gel electrophoresis. For 1% w/v agarose mini gels, 0.4g of agarose was dissolved in 40 ml of 0.5% v/v TBE or TAE (table 2.5) containing 4µl of x10,000 SYBR safe DNA stain (Invitrogen) and heated in a microwave until dissolved. The solution was poured into the casting tray and the comb positioned. Once solidified the casting ends and comb were removed and following loading alongside a 1kb DNA ladder (NEB), gels were run at 90kV, 200mA for one hour. All gels were visualised using a Syngene INGenius ultraviolet trans-illuminator. When smaller fragments required separation, the DNA ladder was substituted for one with smaller increments which ranged from 800bp upto a 1kb ladder all purchased from NEB.

2.2.2.3 DNA quantification

DNA concentrations were determined using a Jenway Genova Nano, Micro-volume spectrophotometer using a sample size of 1.5µl and a corresponding blank, the concentration was deduced using equation 1 and sample purity was assessed using equation 2.

$$\text{Concentration } (\mu g/ml) = (A_{260} - A_{320}) \times 50 \times \text{dilution factor}$$

Equation 1. Calculating DNA concentrations – Barbas et al., 2007.

$$\text{Sample purity} = A_{260}/A_{280}$$

Equation 2. Calculating the DNA purity ratio – Barbas et al., 2007.

Following PCR reactions DNA was quantified using band intensity examination following agarose gel analysis, by comparing the intensity of the product alongside the intensity of a DNA ladder with a known concentration. However, the spectrophotometer reading was the preferred method and was used for calculating ligation reactions.

2.2.2.4 Restriction digest and ligation of targets and vector

All restriction digests were carried out following manufacturer recommendations. PCR products were subject to mini-prep prior to digestion and restriction digests were carried out sequentially in the recommended buffer for each enzyme. Extended restriction digests of both the PCR product and the pET47 vector were performed at 37^{o}C, overnight with gentle agitation (50rpm). The enzymes were then inhibited by heat treatment following manufacturer instructions. Following restriction digestion of both vector and the target gene, samples were subjected to mini-prep and the DNA concentrations obtained. Ligations reactions were carried out using molar ratios of 1:3, 1:5 and 1:10, vector to insert with the amount of vector not exceeding 50ng per reaction. In some cases the amount of vector was taken down to 5ng. Ligation reactions were carried out at 19^{o}C and incubated for 2-4 hours. Following deactivation of the ligase following manufacturer instructions the ligated mixture was split into two samples. The first sample was subsequently used in transformations and the second was subjected to a further digestion using the SalI restriction enzyme, whereby the corresponding restriction site is lost should the vector have been successfully digested by both enzymes. The final digest was incubated at 37^{o}C overnight, deactivated and used to transform *E. coli* cells.

2.2.2.5 Competency and transformation

Where commercial, competent-ready cells were not used, cells were prepared using the calcium chloride method (Mandel & Higa, 1970). Either expression or maintenance strain *E. coli* cells were streaked on to sterile LB plates and following overnight incubation a single colony was transferred to 5 ml of sterile LB and incubated overnight at 37°C. 1 ml of the overnight culture was subsequently transferred to 50 ml fresh LB and incubated as above until an optical density of 0.3 OD_{600} was achieved. The culture was chilled on ice and centrifuged at 3300 x g for ten minutes to pellet the cells. The medium was discarded and the cell pellet was re-suspended in 10 ml of sterile, cold 0.1M $CaCl_2$. The suspension of cells was kept on ice for 30 minutes and the cells pelleted by centrifugation as above. The final cell pellet was re-suspended in 1 ml of sterile, ice cold $CaCl_2$ containing 15% v/v glycerol, aliquoted and stored ready for use at -80°C.

Transformation reaction	Description
Transformation 1	Competent cells + unaltered ligation reaction
Transformation 2	Competent cells + digested ligation reaction
Competency control	Competent cells + unaltered vector
Negative control 1	Competent cells with no DNA
Negative control 2	Competent cells with vector digested with a single restriction enzyme (no ligation)
Negative control 3	Competent cells with vector digested with both restriction enzymes (no ligation)

Table 2.9 - Summary of transformations. Transformations were carried out alongside a series of controls to identify any putatively positive colonies contained the gene of interest. Transformation reactions were carried out in triplicate whereby the amount of plasmid DNA used was varied between 5-50ng.

Transformations were carried out using 50µl of competent cells and various amounts of plasmid DNA in the region of 5-50ng, alongside appropriate controls described in table 2.9. Reactions were incubated on ice for 15 minutes then subjected to heat shock for 45 seconds at 42°C before being returned to ice for 5 minutes. 450µl of LB or super optimal broth (SOC) (table 2.5) was added to the cells and they were allowed to recover at 37°C with gentle shaking

for one hour. The entire volume of cells was then spread onto LB agar with appropriate antibiotic and incubated overnight at 37°C.

2.2.2.6 Restriction analysis and extraction for storage

Following successful transformations, several colonies were selected and individually cultured overnight in 5 ml of sterile lysogeny broth (LB) at 37°C with shaking. Plasmid DNA was then extracted using a QIAprep spin miniprep kit (Qiagen) as per manufacturer guidelines and a sample of the DNA was subjected to restriction digest analysis. Restriction analysis was performed using the same two restriction enzymes, BamHI and NotI. Restriction digests were performed sequentially as previously described (section 2.2.2.4) in line with manufacturer's instructions and the resulting digest analysed by agarose gel electrophoresis (section 2.2.2.2). Those transformations that showed two bands one at a similar size to the pET47 vector and one of a size representable of the gene insert were deemed successful.

DNA from positive transformants was sent to Source Bioscience (Rochdale) for sanger sequencing using standard T7 promoter and terminator primers. The returned sequences were used for nucleotide BLAST (Altschul et al., 1990) searching and the reading frame checked prior to transformation into expression strains.

2.2.3 Protein overexpression

2.2.3.1 Recombinant protein expression by IPTG

Constructs were used to transform suitable expression strains of *E. coli*, as previously described and stocks of transformed cells were stored at -80°C. Freezer stocks were cultured on to LB-agar plates with suitable antibiotic as and when required and incubated overnight; a single colony from the plate was transferred into 10 ml of sterile LB containing appropriate antibiotics and incubated at 37°C overnight with shaking. 1 ml of this overnight culture was then transferred to either 50 ml or 500 ml of LB, again with appropriate antibiotics. The OD_{600} was monitored until the culture reached 0.6 where a small sample was removed for SDS-PAGE analysis, before addition of IPTG to a final concentration of 1mM. Cultures were incubated at 37°C for 4 hours with vigorous shaking or at 15°C overnight. A small sample of the culture was also taken post IPTG induction for further analysis.

2.2.3.2 Recombinant protein expression by auto-induction

Auto-induction media was prepared as indicated in table 2.5. The tryptone and yeast extract, core media was prepared and autoclaved prior to addition of NPS, 5052 and appropriate antibiotic, all of which had been filter sterilised through a 0.45µM unit. As with the IPTG controlled protocol, the auto-induction media was inoculated with 1 ml of overnight culture and incubated for 20-24 hours at 37ºC with vigorous shaking at 200rpm.

2.2.3.3 SDS-PAGE and Western blot analysis

Samples were taken throughout protein production and purification and are further described in table 2.10. Any samples containing guanidine-HCl were desalted prior to SDS-PAGE analysis due to the unfavourable precipitation reaction between the denaturant and SDS. Protein from these samples was obtained by ethanol precipitation and resuspended in either 8M urea or a simple salt buffer. Estimates of protein concentration were made prior to SDS-PAGE analysis and this was achieved either by the deduction of the absorbance at 280 nm or by Bradford protein assay (Bradford, 1976). Following the determination of the protein concentration, samples for SDS-PAGE analysis were prepared using x4 LDS loading buffer (NuPAGE), DTT and deionised water. Sample volumes for analysis varied from 10 to 30 µl dependent upon the gel cast and contained DTT as a reducing agent when required.

Sample	Abbreviation	Description
Pre-induction	---	A sample of the expression culture was taken prior to the addition of IPTG and was harvested as described in section 2.4.
Post-induction	---	A sample was taken following the full expression period.
Total cell lysate	Lysate	A sample was retained following cell lysis.
Soluble	Sol	A sample of the soluble fraction was acquired by centrifugation following cell lysis.

Inclusion body washes	---	Samples of both the supernatant and pellet were retained from each wash during the inclusion body preparation procedure.
Insoluble	Insol	The insoluble matter was subjected to its corresponding denaturant and a sample of the insoluble cellular proteins was obtained from the supernatant following centrifugation. .
Immobilised metal affinity chromatography (IMAC) flow through	FT	A sample of IMAC flow through following loading of the protein sample.
IMAC wash peaks	---	Samples taken from any peaks in absorbance at 280 nm during a low imidazole wash.
IMAC eluate	Eluate	A sample of the IMAC eluate fractions were retained for analysis and positive fractions were pooled.
Ion exchange chromatography (IEC) flow through	FT	A sample of IEC flow through following loading of the protein sample.
IEC eluate	Eluate	A sample of the IEC eluate fractions were retained for analysis and positive fractions were pooled.
Post dialysis		Samples taken following the concentrating and dialysis of fractions to determine protein condition.
Prior to structural experiments Small angle X-ray scattering (SAXS), circular dichroism (CD) and crystallography		SDS-PAGE carried out prior to all experiments to determine protein condition
Post SAXS		Samples analysed following SAXS experiments to determine protein condition

Table 2.10 - A summary of the samples retained for SDS-PAGE analysis. Samples for SDS-PAGE analysis were taken throughout all experimental stages in order to confirm, purity and degradation. Where required samples were reduced using DTT and boiled for 10 minutes at 70^{o}C prior to loading.

2.2.3.4 Casting of gels and running conditions

Where 4-12% bis/tris, pre-cast gels were not utilized (Invitrogen, Paisley, UK), 10% gels were cast using the PerfectBlue Dual Gel Twin S gel tank system (Peqlab, Salisbury Green, UK). A 10 ml resolving gel was cast using of 6.75 ml ultrapure water, 3.75 ml 1.5M Tris pH8.8, 3.75 ml bis/tris acrylamide (40% v/v), 75 µL of SDS (20% w/v), 150 µl APS (10% w/v) and 15 µl TEMED (Final TEMED concentration 0.9% v/v). The stacking gel was prepared with 3.13 ml ultrapure water, 1.25 ml Tris pH 6.8, 0.62 ml of bis/tris acrylamide (40% v/v), 25 µl of SDS (20% w/v), 50µl APS (10% w/v) and 5µl TEMED (Final TEMED concentration 0.9% v/v). Novex sharp pre-stained or un-stained protein standards were used at a volume of 7.5µl as a marker and gels were stained for an hour in coomassie blue R-250 before being subjected to de-stain. SDS-PAGE was run in MES SDS running buffer (as described in table 2.5) at 180 V, 100mA for 40-50 minutes, or until the loading dye had reached the bottom of the gel. When improved resolution was required SDS-PAGE was run more slowly at 100mA and the tank system was either ice packed or water cooled if required.

2.2.3.5 Western blot analysis

All immunoblotting was undertaken using an XCell II blot module (Invitrogen) on ice. Following SDS-PAGE, gels were prepared for transfer by removal of the wells and foot. PVDF transfer membrane was activated by methanol and wet transfer was achieved at 30 V for 90 minutes using transfer buffer as described in table 2.5. The blotting module was arranged as shown in figure 2.4 where two sponges and two sheets of filter paper are soaked in transfer buffer and stacked with the gel and the membrane in the middle with filter paper and a sponge at each side. This stack was then positioned so that the gel was underneath the PVDF and closest to the cathode.

Following blocking in 2% w/v milk powder made up in TBS (table 2.5), appropriate primary and secondary antibodies were used according to manufacturer instructions as described in table 2.11. Most primary antibodies were made up at a ratio of 1:1000 in TBS-Tween and the membrane incubated for 1 hour at 50 rpm/4^{o}C. Following the primary incubation the membranes were washed with 5 ml of TBS-Tween for 5 mins with rapid shaking (150 rpm/room temperature) three consecutive times prior to the addition of the secondary antibody at a ratio of 1:5000 made up in TBS-Tween. Secondary incubations were carried out for at least 1 hour at 50 rpm/4^{o}C, prior to washing with TBS-Tween three times as previously described.

Membranes were then visualised using a LI-COR Odyssey infrared imaging device (LI-COR Biosciences) at 680 nm.

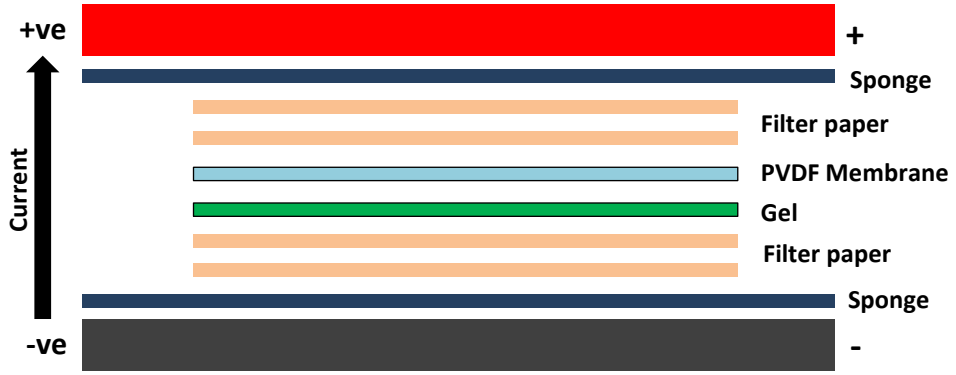


Figure 2.4 – Western blot module packing. Typical arrangement within the western blotting module. The gel is positioned underneath the PVDF membrane and closest to the cathode.

Antibody	Supplier	Description	Ratio (In TBS-Tween)
Monoclonal mouse anti-his	Invitrogen (MA1-21315)	Primary antibody used to detect his-tagged proteins	1:1000
Monoclonal mouse anti-human factor H	Abcam (ab118820)	Primary antibody to detect human factor H	1:1000
Polyclonal goat anti-mouse AlexaFluor 680	Invitrogen (A-21057)	Secondary antibody used for fluorescent detection at 680nm.	1:5000

Table 2.11 - Antibodies used for western blotting – All antibodies were used subject to manufacturer guidance. Ratios were optimised for each particular western blot and antibody solutions were never re-used.

2.2.4 Protein extraction and preparation of OMP-A like proteins

2.2.4.1 Cell harvesting and lysis

E. coli expression cells were pelleted by low speed centrifugation (10,000 x g) at 4ºC, the supernatant was discarded and the cells re-suspended in lysis buffer (table 2.5) at one twentieth of the original culture volume and kept on ice. The suspension was separated into 15 ml falcon tubes and the cells disrupted by pulsed sonication on ice (40% amplitude, 5 seconds on, 10 seconds off for a total of five minutes). The disrupted suspension was centrifuged at 20,000 x g for 1 hour or until a firm pellet was formed. The supernatant or soluble fraction was retained for analysis and the insoluble pellet kept for inclusion body preparation.

2.2.4.2 Inclusion body preparation and solubilisation

In order to obtain a high level of pure recombinant protein, the insoluble expression material was subjected to a DNase incubation and a multi-cycle, wash step. The insoluble material was resuspended in 0.3M NaCl, 50mM Tris base, pH7 and briefly pulsed sonicated at an amplitude of 40% for a total length of 30 seconds. DNase1 was added at 1 µg/ml and the solution incubated at 4ºC for 4-16 hours. Following incubation, samples were centrifuged until all insoluble material was pelleted and weighed prior to re-suspension in wash buffer 1 (Table 2.5) at 10 ml per gram. The suspension was briefly sonicated as previously described. Following this, the suspension was subjected to low speed centrifugation (~8000 x g) until a firm pellet was obtained. This process was repeated twice with inclusion wash buffer 1, twice with wash buffer 2, and twice with wash buffer 3. The inclusion body was then solubilized in denaturing buffer overnight at room temperature with shaking.

2.2.5 Protein purification and refolding of OMP-A like proteins

2.2.5.1 Immobilised Metal Affinity Chromatography (IMAC)

Solubilized inclusion bodies were clarified by centrifugation (10,000 x g) or by filtration prior to all chromatographic steps. IMAC was carried out using a 5 ml HisTrap HP Ni-NTA column (Qiagen and GE Healthcare) and assisted by an Akta Prime FPLC system (GE Healthcare). The column was equilibrated with ten column volumes of appropriate FPLC binding buffer (table 2.4), prior to protein sample loading. The flow rate of protein samples was varied for optimal binding but was kept below 1 ml/min. Following protein binding, equilibration buffer was fed through the column until a steady absorbance baseline was reached (280nm), followed

by 10 column volumes of FPLC wash buffer to remove any non-specifically bound proteins. The target protein was eluted by the competitive addition of imidazole to the column, in the form of an increasing gradient to 0.3M imidazole over a total volume of 50 ml. Flowthrough during protein loading and the eluate following wash steps was collected in 2 ml fractions and retained for SDS-PAGE analysis. The final eluate was collected in either 0.5 ml or 1 ml fractions and retained at 4^{o}C for further use.

2.2.5.2 Ion Exchange Chromatography

Inclusion bodies were solubilised in low salt denaturing buffer (less than 20mM NaCl) and the cleared sample was loaded on to a 5 ml anion exchanger column (HiTrap Q – GE Healthcare), which had previously been equilibrated with 5 column volumes of IEC binding buffer (Table 2.4). The sample was then eluted by increasing NaCl concentration over a variety of gradients ranging from 40 to 100 ml at a flow rate of 1 to 3 ml/min.

2.2.5.3 Bench top refolding

Urea was removed from denatured protein samples by rapid dilution, protein samples were added to re-naturing buffer (at a volume 50 times larger than the original denatured protein volume and at an overall concentration of below 0.1 mg/ml) drop-wise, over several minutes. The diluted protein solution was left stirring overnight, prior to concentrating in a 10,000 MWCO Amicon centrifugal device or where large volumes were generated by pumping onto a Ni-NTA column using a FPLC.

2.2.5.4 On-column refolding

On column-refolding was carried out by both IMAC and IEC, once the protein sample was bound to the target column the denaturing agent, in most cases urea, was removed by a shallow gradient from denaturing buffer to native buffer with appropriate detergent. Refolding gradients were run between 50 to 200 ml at slow flow rates, below 0.8 ml/min.

2.2.5.5 Removal of the poly-histidine-tag

The poly-histidine, purification tag of BAPKO_0422 was removed prior to some crystallisation trials and SAXS experiments. The protease, HRV-3C was employed for cleavage of the recognition site LEVLFQ↓GP. Protein samples were subjected to digests ranging from 1:100 to 1:2 HRV-3C to protein and kept either at room temperature overnight or at 4^{o}C for 48 hours with gentle agitation. IMAC was performed using a Ni-NTA 5 ml column to remove the protease, polyhistidine tag and un-digested protein.

2.2.5.6 Concentration and dialysis of protein samples

Following both purification and tag removal, protein samples were adjusted to the desired concentration by the use of an Amicon centrifugal concentrator device 10,000 (MWCO). Following any concentrating procedure extensive dialysis was carried out using a Slide-A-Lyzer 10k MWCO dialysis cassette (Thermo Scientific). Refrigerated protein dialysis was carried out over a minimum of three days with two buffer changes on day one. Dialysis was performed in a 500 ml beaker filled with 500 ml of the intended end buffer. Following dialysis, protein samples were used immediately for structural experiments or stored at $^{-}$20^{o}C in 10% v/v glycerol.

2.2.5.7 UV absorbance spectrum of refolded protein and quantification

Following protein purification and refolding, the UV absorbance spectrum was obtained for each sample. Samples were blanked against their corresponding buffers and scanned from 220 to 340 nm using a Spectrostar Nano UV spectrophotometer (BMG Labtech) in order to assess any DNA contamination (figure 2.5). Samples were diluted where required in order to achieve a reading between 0.1-1. Reading were taken using 300µL of protein solution in a disposable plastic UV cuvette (Eppendorf). Protein concentrations were determined using theoretical extinction co-efficients, calculated using the online tool ProtParam by ExPASy and the Beer-Lambert Law (Equation 3) alongside the reading at 280 nm minus the reading at 340 nm.

$$A = \varepsilon c l$$

Equation 3 - The Beer-Lambert Law. Where A, absorbance; ε, molar extinction coefficient; c, concentration; l, pathlength.

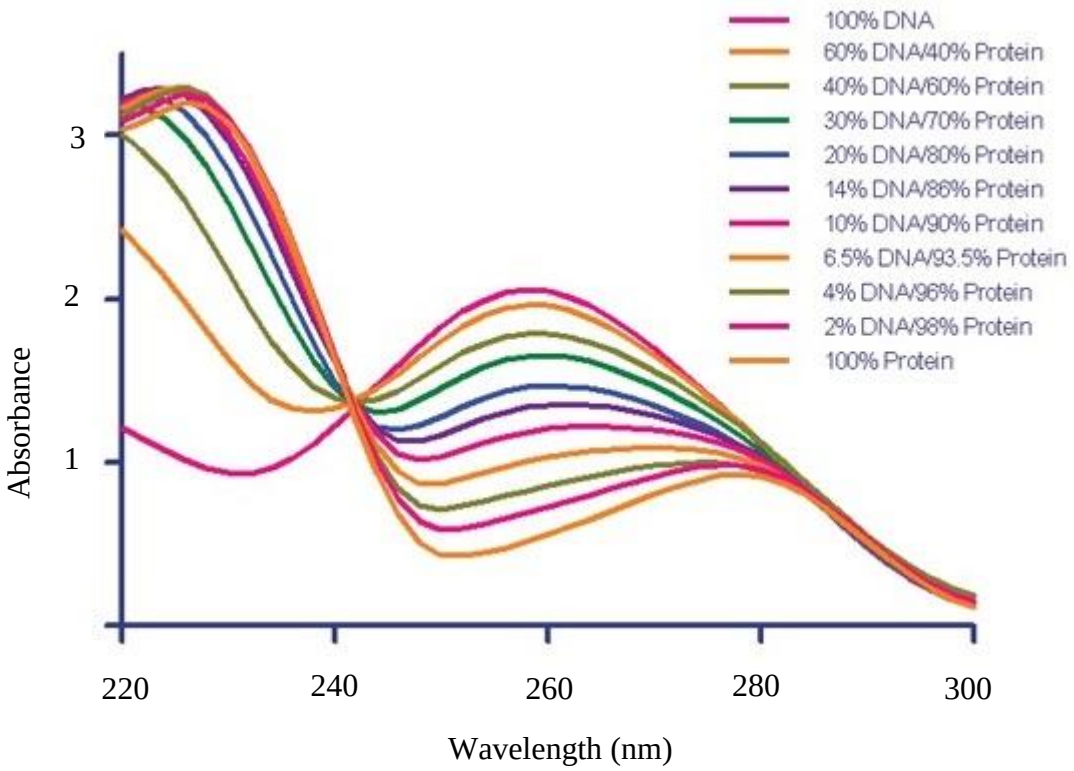


Figure 2.5. UV spectral scans of mixed samples of protein and dsDNA. Samples of purified dsDNA (Herring sperm dsDNA) and protein (BSA) at different ratios and their corresponding UV spectrum. Figure adapted from application notes from Biotek.com.

2.2.5.8 Size Exclusion Chromatography (SEC)

SEC or gel filtration chromatography was used as a final polishing step and to remove any aggregation from the sample prior to structural experiments whilst allowing an estimation to be made of the size of the protein. Prior to all gel filtration runs the void volume of the column was determined using blue dextran and a calibration curve prepared using the known protein standards shown in table 2.12 and the calibration curve shown in figure 2.6. Protein standards were purchased from Sigma and reconstituted into 0.3M NaCl, 50mM tris base at pH8 prior to SEC.

Protein Standard	Molecular weight (kDa)	Elution volume (ml)
Vitamin B12	1.35	17.55
Ribonuclease A	13.7	13.43
Ovalbumin	45	11.3
Bovine serum albumin	66.5	10.19

Table 2.12 - SEC calibration proteins – A summary of the proteins standards and their molecular weight used to construct a calibration curve. All proteins were prepared in SEC buffer (0.3M NaCl, 50mM tris, pH8).

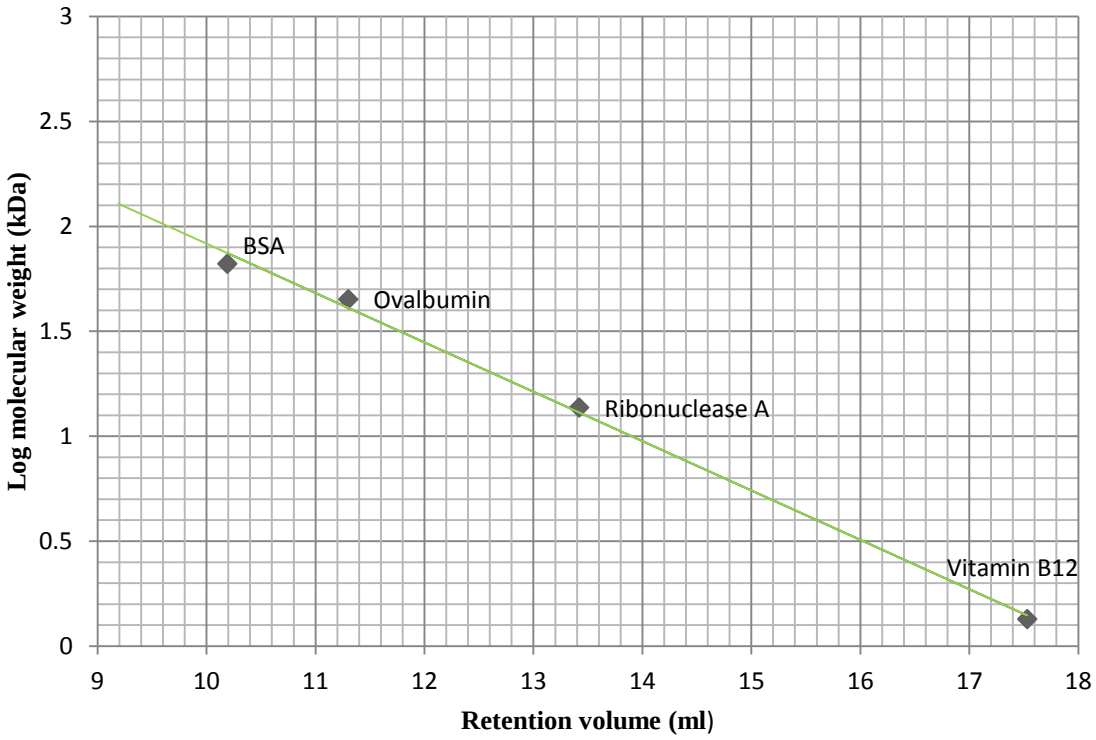


Figure 2.6 - Size exclusion calibration curve of known protein standards - Protein standards at 1 mg/ml were prepared in SEC buffer (components shown in table 2.4) filtered and 200µl injected into the top of the pre-equilibrated column. The flow rate was set at 1 ml/min and the absorbance monitored at 280nm.

Protein samples were subjected to SEC to determine an approximation of the molecular weight and to attain a monodisperse sample for further experiments. SEC was carried out using a Superdex 75 10/300 GL column (GE Healthcare) which was equilibrated with appropriate native buffers in most cases for OMPs this was 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO at pH8. Protein samples were concentrated to a volume of approximately 100µl with the

concentration not exceeding 10 mg/ml and injected onto the top of the column with a 1 ml disposable syringe.

2.2.6 Characterisation of OMP-A like proteins

2.2.6.1 Gel shift assays for *Borrelia* OMPs

Gel shift assays were conducted according to published methods (Laemmli, 1970, Arora et al., 2000, Tamm et al., 2004, Verhoeven et al., 2013). Samples of *Borrelia* OMPs were refolded into 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8 and the assay undertaken using both native and SDS-PAGE. Two samples for each *Borrelia* OMP were prepared with one being subjected to a 10 minute incubation at 80^{o}C and the second remaining at room temperature. Further gel shift assays were performed using varying concentrations of urea to assess whether protein migration during SDS-PAGE was affected by unfolding. *Borrelia* OMP proteins were prepared in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO with varying levels of urea from 8M to 1M and subjected to SDS-PAGE according to published methods.

2.2.6.2 Phase partitioning of *Borrelia* OMPs

Phase partitioning experiments were conducted according to previous published methods (Bordier, 1981). Proteins, haemoglobin, lysozyme and a voltage-dependent anion channel (VDAC) from *Arabidopsis thaliana* were selected as control proteins and were purchased from Sigma-Aldrich with the exception of the VDAC which was kindly supplied by Dr George Psakis and had been expressed, purified and refolded according to published methods (Mertins et al., 2012).

The cloud point of a detergent in solution is the temperature at which it separates into two distinct phases and is normally observed when the detergent is close to or above it's critical micelle concentration (CMC). As the micelles begin to aggregate two phases appear, one which normally appears cloudy caused by the micelle aggregation the second phase becomes depleted of detergent. This feature of detergents can be exploited in order to separate proteins which interact with detergents from those which do not. In order to do this a detergent with a low cloud point temperature is required with Triton X-114 most often being selected as it's cloud point is approximately 23^{o}C. However, this can be affected by the addition of salts or polymers.

For phase partitioning experiments all proteins solutions were prepared at 1 mg/ml in 10mM Tris-HCl pH7.4, 150mM NaCl, 1% v/v Triton X-114 0^{o}C. A sucrose cushion consisting of 6% w/v sucrose, 10mM tris/HCl, pH 7.4, 150mM NaCl with 0.06% v/v Triton X-114 was prepared in an Eppendorf at a total volume of 200µl. On top of the sucrose cushion 50µl of ice cold protein solution was overlaid and incubated at 30^{o}C for three minutes prior to centrifugation at 300 x g for three minutes. The aqueous upper layer was removed (approximately 70µl) and supplemented with additional Triton X-114 to a concentration of 0.5% v/v. Following gentle agitation and a further three minute incubation at 0^{o}C the aqueous phase was returned to the same sucrose cushion and incubated for 3 minutes at 30^{o}C prior to centrifugation at 300 x g for three minutes in order to pellet the detergent phase further. Again the upper aqueous phase was removed and transferred to a fresh Eppendorf and was further supplemented with Triton X-114 to a concentration of 2% v/v and incubated at 0^{o}C for 3 minutes. Following the incubation the aqueous phase was incubated at 30^{o}C for 3 minutes and subjected to a final round of centrifugation at 300 x g for 3 minutes. The aqueous phase following centrifugation was retained for analysis and the pellet discarded. The detergent phase from below the sucrose cushion was removed by aspirating 10µl from the very bottom of the Eppendorf and retained for SDS-PAGE analysis.

This procedure was carried out for all three control proteins alongside the four proposed *Borrelia* OMPs. SDS-PAGE analysis of the resulting phases was performed as previously described in section 2.2.3.3.

2.2.6.3 Affinity ligand binding assays for *Borrelia* OMPs

Purified and re-folded OMPs (20-50µg) were subjected to non-reducing SDS-PAGE alongside human factor H (20µg) and bovine serum albumin (20µg) as positive and negative controls respectively. The gel was immunoblotted, the membrane trimmed and then blocked in 5% w/v milk powder made up in TBS. Following three, 5 minute washes with TBS-Tween (0.05% v/v) the membrane was incubated with human factor H, at concentrations between 5 to 88 µg/ml made up in TBS and left at 4^{o}C, overnight with gentle agitation. The membrane was again washed as above with TBS-Tween, before incubation with the primary antibody for one hour (1:1000 mouse monoclonal anti-human factor H – Abcam). Following further washing the membrane was subjected to the secondary antibody (1:5000 of goat anti-mouse polyclonal

alexafluor 680 – Abcam) incubated for one hour , washed and visualised on a LI-COR Odyssey infrared imaging device at 680nm.

2.2.6.4 Immobilised heparin binding assays for *Borrelia* BAPKO_0422

Purified and re-folded recombinant BAPKO_0422 was pumped onto a pre-equilibrated (x5 column volumes of 0.1M NaCl, 50mM tris, 0.1% v/v LDAO at pH8) GE Healthcare HiTrap HP-5 ml Heparin column using an Akta Prime chromatography system. A total volume of 3 ml of BAPKO_0422 in solution (BAPKO_0422 at 0.5 mg/ml in 0.1M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8) was loaded onto the column at a flow rate of 1 ml/min. The flow through was collected and retained for SDS-PAGE analysis. The column was further washed using 10 column volumes of the equilibration buffer with all resulting flowthrough collected and retained. Following extensive washing the resulting protein was eluted by means of a 50 ml gradient from equilibration increasing upto 1M NaCl, 50mM tris base, 0.1% v/v LDAO at pH8, 1 ml fractions were collected and used for SDS-PAGE analysis.

2.2.7 Protein overexpression and preparation of superoxide dismutase A

2.2.7.1 Bacterial Strains, Plasmids, Growth Conditions and Protein Expression

In order to perform SOD activity assays, the expression and refolding methods which have previously been described in sections 2.2.2 and 2.2.3 were altered and are summarised below. The *bb0153* gene was codon optimised for expression in *E. coli*, synthetically produced and sub-cloned into pET-47b(+) (EMD Millipore). The resulting construct was transformed into T7 Express Competent *E. coli* (New England BioLabs). Transformed cells were grown in LB media at 37^{o}C in an orbital incubator and expression induced for 4 hours by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a working concentration of 1mM. Induced cells were harvested by centrifugation and lysed on ice by pulsed sonication in 0.3M NaCl, 50mM Tris base, pH8. The soluble fraction was removed by centrifugation (20,000 x g for 30mins at 4^{o}C) and the pellet solubilised in 8M urea, 0.3M NaCl, 50mM Tris base, pH8 before clarification by centrifugation.

2.2.7.2 Production of apo-protein

Existing metal ions were removed by utilising the properties of nitrilotriacetic acid (NTA) by the means of a stripped 5 ml Ni-NTA column (Qiagen). The column was stripped of Ni^{2+} as advised within the user manual and the recombinant protein sample was passed twice through the column at a flow rate of 0.5 ml/min using an AKTA Prime FPLC system.

2.2.7.3 Addition of metal co-factors

Following the removal of bound metal ions by NTA the pH of the denatured protein solution was decreased to pH4 by slow addition of dilute hydrochloric acid. $MnCl_2$ and $FeCl_2$ were added to samples of recombinant SodA to a final concentration of 1mM. A third sample of the apo-protein was retained with no additions made. Recombinant protein samples were incubated at 4^{o}C overnight before the pH was increased back to pH8 by slow addition of dilute sodium hydroxide.

2.2.7.4 On-column refolding and purification

Samples were subjected to centrifugation at 16,000 x g for 30 minutes prior to purification using a Hi Trap Q HP column (Amersham Bioscience). The column was equilibrated with 10 column volumes of binding buffer (8M urea, 50mM Tris Base, 0.3M NaCl, pH8) and each recombinant protein sample loaded at 0.5 ml/min. Following loading the protein was refolded into 1M urea, 50mM Tris base, 0.3M NaCl, 0.05% v/v LDAO, pH8, by a linear gradient over 60 ml at a flow rate of 0.5 ml/min. Refolded SodA was then subjected to 10 column volumes of wash buffer, 1M urea, 50mM tris base, 0.3M NaCl, 50mM imidazole, pH8 before being eluted by linear gradient of 50mM tris base, 0.3M NaCl, 0.3M imidazole, 0.05% v/v LDAO at pH8. This process was repeated for each SodA sample, fractions containing the refolded protein were respectively pooled and subjected to SDS-PAGE analysis. Mn-SodA, Fe-SodA and apo-SodA were concentrated and buffers exchanged (0.3M NaCl, 50mM Tris base, 0.05% v/v LDAO, pH8) using a 10NMWL Amicon Ultra-15 centrifugal filter unit (Merck Millipore), centrifuged at 16,000 x g for 20minutes at room temperature and their concentration ascertained by spectroscopy and Bradford analysis (Bradford, 1976).

2.2.7.5 Assaying SOD activity

In solution superoxide dismutase activity was accessed using a SOD determination kit from Sigma. The assay was performed according to the manufacturer instructions with additional negative controls included to assess any possible interference from the utilized buffers. Although there are several methods in order to determine SOD activity both in solution or from a protein sample within a gel following electrophoresis separation the in solution Sigma-Aldrich SOD determination kit was selected as it would allow a range of additional controls to be added. The assay kit measures SOD activity using WST-1 (2-(4-Iodophenyl)- 3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H tetrazolium monosodium salt) a highly water soluble tetrazolium salt that produces a water soluble formazan dye following reduction with a superoxide anion. The rate of this reduction by the superoxide anion can be described as linearly related to xanthine oxidase activity as shown in figure 2.7 which is inhibited by SOD allowing inhibition rates to be monitored by spectrophotometry via this colorimetric assay.

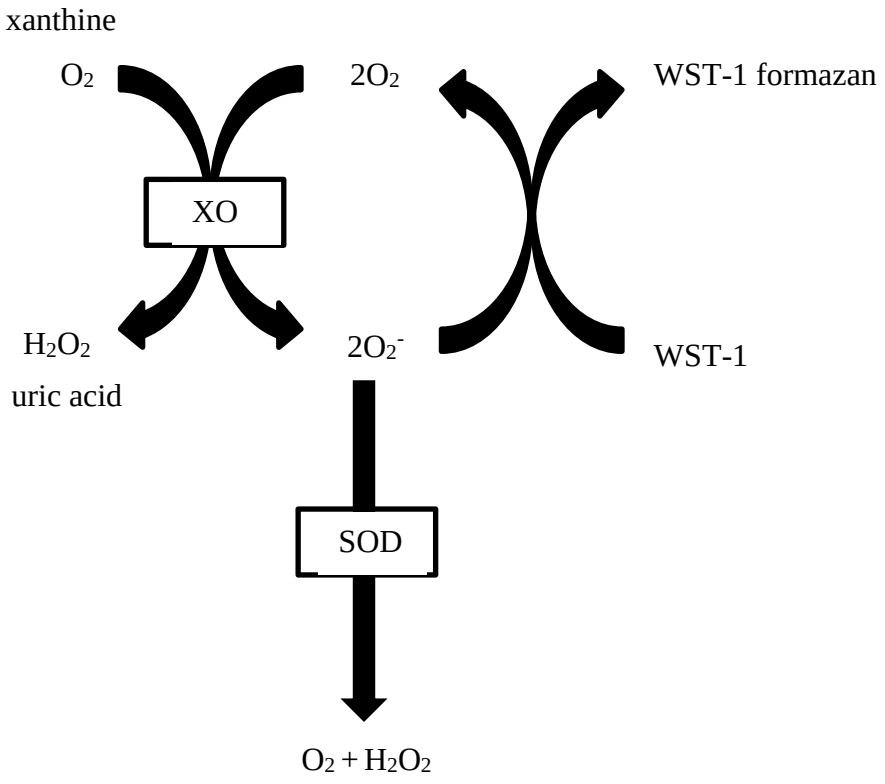


Figure 2.7. Reaction scheme for SOD activity determination. Indirect determination of SOD activity (SOD determination kit – Sigma 19160). The SOD assay kit assays SOD activity using a water soluble tetrazolium monosodium salt – WST-1. WST-1 produces a water soluble formazan dye upon reduction with superoxide anions. As the rate of reduction with O_2 is linearly related to xanthine oxidase activity which is inhibited by SOD. Measuring the decrease in colour development at 440 nm allows the inhibition activity to be quantified.

The kit was used to compare activity rates of recombinant Mn-SOD, Fe-SOD and apo-SOD and run in 96 well plates with a total well volume of 240 µl in quadruplets. Changes in absorption were measured at 440 nm using a BMG Labtech SPECTROstar Nano microplate-reader. The assay was prepared according to manufacturer instructions (Sigma Aldrich – 19160 SOD determination kit) and as presented in table 2.13. The enzyme working solution was added to each well last and using a multichannel pipette as this initiated the reaction. Following set up the microplate was mixed by shaking (80rpm) in a SPECTROstar microplate reader (BMG Labtech) and incubated at 37^{o}C. The absorbance was measured at 450 nm every minute for a total of 20 minutes.

Reaction	SodA stock protein (µl)	ddH_2O (µl)	WST working solution (µl)	Enzyme working solution (µl)	Dilution buffer (µl)	SodA buffer (µl)
Blank 1	-	20	200	20	-	-
Blank 2	-	13.02	200	20	-	6.98
Blank 3	-	13.33	200	20	-	6.67
Blank 4	-	12.77	200	20	-	7.23
Blank 5	-	20	200	-	20	-
Blank 6	6.98	13.02	200	-	20	-
Blank 7	6.67	13.33	200	-	20	-
Blank 8	7.23	12.77	200	-	20	-
Mn-SodA	6.98	13.02	200	20	-	-
Fe-SodA	6.67	13.33	200	20	-	-
Apo-SodA	7.23	12.77	200	20	-	-

Table 2.13. Reaction volumes for assaying SOD activity. Reactions were set up according to manufacturer instructions (Sigma Aldrich – 19160 SOD determination kit). Blanks 2,3 and 4 were additional controls to ensure SodA buffers (0.3M NaCl, 50mM tris, pH8) did not interfere with the reaction. SodA stock protein volumes were calculated in order to achieve a well concentration of 100mg/ml for each reaction. The enzyme working solution was added last and using a multi-channel pipette as this initiated the reaction.

2.2.8 Structural Experiments of OMP-A like proteins and SodA

2.2.8.1 Circular Dichroism - Sample preparation, data collection and analysis

Circular dichroism experiments were carried out on a Jasco J-810 spectropolarimeter with the kind assistance of Dr Andrew Leech (University of York). Air and buffer blank background data was acquired and experimental measurements taken over a wavelength range of 185-260 nm. Experimental data was obtained over five individual scans and averaged. A full summary of the experimental parameters are shown in table 2.14. Protein samples were prepared in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO at pH8 for OMPs and the same buffer but without the detergent LDAO for SodA. Protein samples were taken from high concentration protein stocks (~10mg/ml) and diluted using the same buffer to between 0.5 and 0.05mg/ml.

Data acquired was analysed using DichroWeb (Whitmore and Wallace, 2004, Whitmore and Wallace, 2008). Algorithms used for deconvolution included CDSSTR (Compton & Johnson, 1986, Manavalan & Johnson, 1987), CONTIN (Provencher & Glockner, 1981, Van Stokkum et al., 1990) and SELCON3 (Sreerema & Woody, 1993, Sreerema et al., 2000) alongside SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011) reference sets, further discussed in table 1.

Experimental temperature	20ºC
Path length	0.2cm
Protein concentration	0.1 mg/ml
File format	JASCO 1.30
Wavelength range and step	185-260 nm in 0.5 nm increments
Analysis algorithm	CDSSTR (Compton & Johnson, 1986, Manavalan & Johnson, 1987)
Reference sets	SP175(short) (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011)
Initial wavelength	260 nm
Final wavelength	185 nm
Input units	Mdeg/theta machine units
Output units	Mean residue ellipticity or delta epsilon
Mean residue weight (Da)*	111.8

Table 2.14. Parameters used during CD data collection and Dichroweb analysis

*Mean residue weight in Daltons is the average molecular weight of all the amino acids within the protein sequence. The value is calculated by dividing the total molecular mass of the polypeptide by the number of residues.

2.2.8.2 Small Angle X-ray Scattering

2.2.8.2.1 Sample preparation and data collection

Purified protein samples for both in-house and synchrotron data collection were prepared in SAXS buffer (0.3M NaCl, 50mM tris base, 0.1% v/v LDAO - table 2.5) and subjected to extensive dialysis as previously described (section 2.2.5.6). Where LDAO was used extra care was taken in the preparation of blank samples. Blanks were obtained directly from the final dialysis buffer and stored at $^{-}20^{o}$C. Control samples of Lysozyme were prepared at various concentrations ranging from 0.5-8mg/ml, both with and without LDAO and data collected for analysis as a control protein. In house scattering was collected on a BRUKER NANOSTAR all experiments were designed to obtain at least 300,000 counts per frame with the system under vacuum conditions. Sample volumes ranged from 60-100µl and held within a quartz glass capillary (diameter 1.5mm). Data for both the blank and the sample was collected in sets of 10 frames with an exposure time of 2400 seconds per frame.

2.2.8.2.2 Data analysis

Following data collection all frames, both blank and sample were integrated by the chi axis using Bruker inbuilt software and an output file (.dat) generated. Integrated frames collected for both the blank and each sample were independently averaged using PRIMUS (Konarev et al., 2003) and then the averaged blank scattering deducted from the averaged sample scattering. This subtracted (.dat) file was then used as the starting point when working with the ATSAS (Svergun, 1992, Franke & Svergun, 2009, Volkov & Svergun, 2003, Kozin & Svergun, 2001, Svergun, 1999) suite of software. The devised pipeline for data analysis is shown in table 2.15.

Stage	Program	Summary	Reference
1	Primus	Integrated scattering data is averaged and subtracted then visually checked for aggregation and consistency. Kratky and Guinier analyses are undertaken and an initial estimation of the radius of gyration may be made.	Konarev et al., 2003
2	GNOM	An indirect transform of the data is carried out and the P(r) and *Rg* can be calculated. The resulting transformed data is saved as a .out file which can be used for ab initio modelling.	Svergun, 1992
3	DAMMIF	Actions ab initio bead modelling. The program is run using the same data set twenty times to give multiple possible models.	Franke & Svergun, 2009
4	DAMSEL	All generated DAMMIF models are input into DAMSEL which superimposes all models and removes the least probable solutions.	Volkov & Svergun, 2003
5	DAMSUP	All models excluding DAMSELS outliers are taken by DAMSUP, the most probable solution is selected and all models are aligned to this.	Volkov & Svergun, 2003
6	DAMAVER	The aligned models are averaged by DAMAVER and a probability map is calculated.	Volkov & Svergun, 2003
7	DAMFILT	The final model is generated by DAMFILT which uses a given cut off value and removes regions of low occupancy.	Volkov & Svergun, 2003

Table 2.15. SAXS data analysis pipeline. The ATSAS suite of programs was used to generate three dimensional models for each protein sample. The analysis process was tested with two control proteins further discussed in section 3.10.4 of the results chapter.

2.2.9 Crystallography of SOD and OMP targets

2.2.9.1 Vapour diffusion crystallography

Vapour diffusion crystallography (Benvenuti & Mangani, 2007, Drenth, 2007) was carried out using 24 well trays and set up in hanging drop format. Crystallisation screens utilized throughout the project are listed in table 2.16. All screens were formulated in house with the exception of the MembPASS screen which was purchased from Jena Bioscience. Well conditions were created using stock chemicals which were centrifuged and filtered using a 0.45µm filter unit before being appropriately stored. Proteins were concentrated between 10 to 15 mg/ml as previously discussed and the drop was composed of 1 µl of concentrated protein and 1 µl of the mother liquor/well solution. Trays were incubated at 20^{o}C and checked every day for the first week and once per week there after using a Leica stereozoom benchtop microscope.

Crystallisation Screen	Manufacturer
Clear strategy screen 1 – CSS1	Molecular Dimensions
Clear strategy screen 2 – CSS2	Molecular Dimensions
Wizard Screen 1	Rigaku
Wizard Screen 2	Rigaku
Wizard Screen 3	Rigaku
Wizard Screen 4	Rigaku
MembPASS screen	Jena Bioscience

Table 2.16. Crystallisation screens used within the project. All screens with the exception of the MembPASS screen were created in house following the manufacturer's formulations.

2.2.9.2 In meso phase crystallography

In meso phase or lipidic cubic phase crystallography (LCP) was used in attempt to crystallise the OmpA-like proteins (Caffrey, M. and Cherezov, 2009). The basis of the method is the utilisation of a stable lipid cubic phase in order to provide a more natural environment for membrane protein crystallisation. Monoolein or MO (1-oleoyl-rac-glycerol) is the lipid of choice for this method although others can be used or combined with MO. In order to reconstitute the protein samples into the lipid environment a gas tight mixing device was created by the joining of two gas tight syringes using a coupling unit, which was constructed by the silver soldering of two syringe nuts together with a custom size needle positioned inside (as shown in figure 2.8).

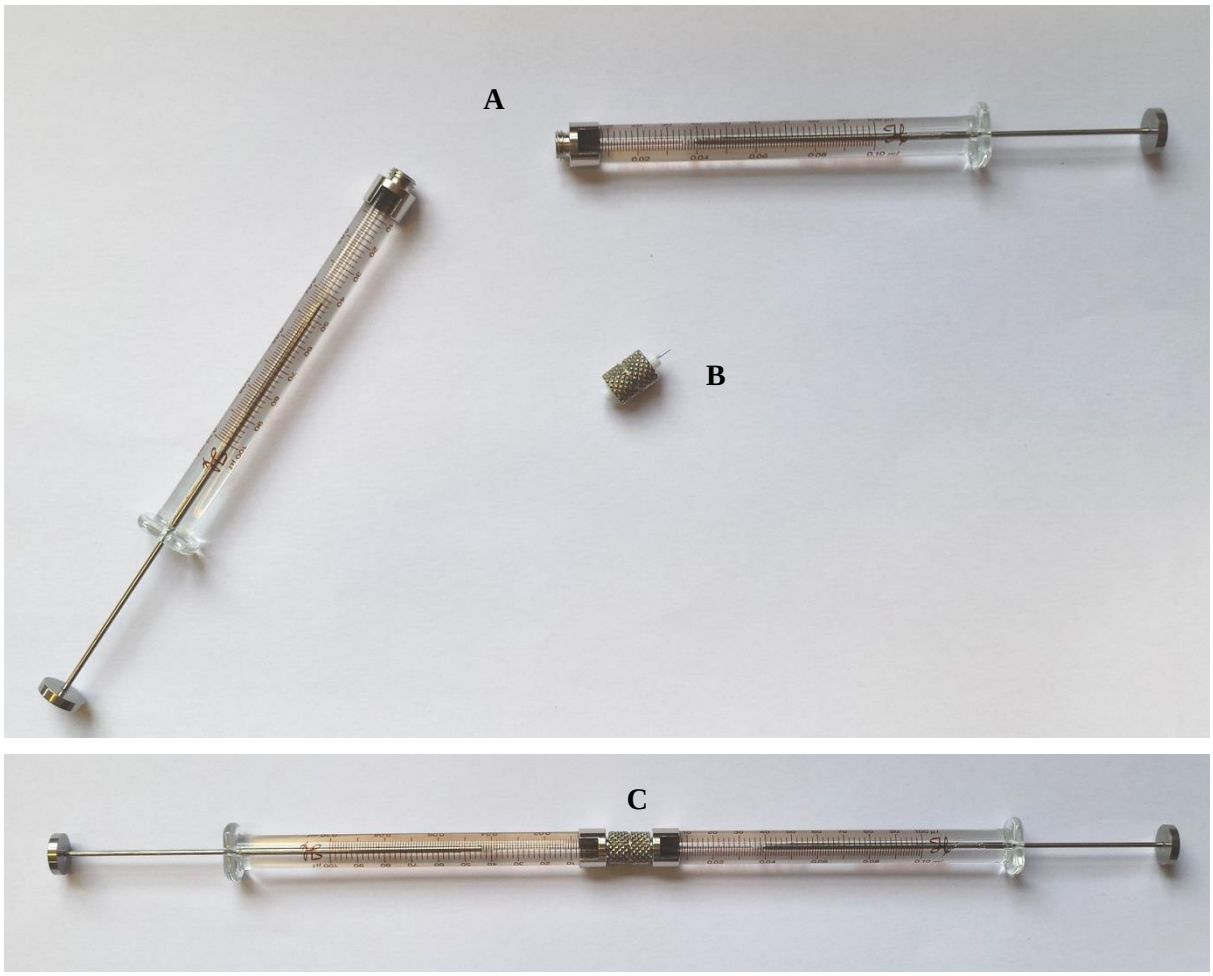


Figure 2.8. Production of the LCP mixing device. **A.** Two Hamilton gas tight 200µl syringes with plastic ferrules removed. **B**. Two R nuts silver soldered together with a custom made 26 gauge needle through the middle held in place with a double plastic ferrule at each end. **C**. Final lipid/protein mixing device.

LCP crystallography was set up as shown in figure 2.9 using 15-50 mg MO which was transferred to a plastic vial and the lipid melted by incubation at 37^{o}C for 10 minutes. The molten lipid was then transferred to a 100µl Hamilton gas tight syringe with an attached syringe coupler, the weight of both the empty and filled syringe was recorded and the total amount of MO transferred calculated. Using a 25 µl gas tight Hamilton syringe with a flat tipped 26 gauge needle the concentrated protein was transferred to a second 100 µl Hamilton gas tight syringe. The amount of protein transferred was calculated in order to achieve a 40% w/v of protein solution within the final mixture for example; 30 mg of lipid was used with 20 µL of concentrated protein. The contents of both syringes were pushed up through the barrel to ensure no air bubbles were trapped and then the syringe containing protein was connected to the MO syringe via the coupling unit. The contents were mixed back and forth through both syringes, for approximately 5 minutes or until a homogenous cubic phase was achieved.

Following mixing the protein/lipid paste was transferred across to a single syringe and the coupling unit replaced with a single 26 gauge needle. The loaded syringe was attached to a Hamilton dispensing unit and 0.5 µl of mixture was ejected into each well of a 96 well LCP screening plate (Swissci). Each well was overlaid with 0.5 µl of precipitant solution and the plate sealed with UPV film for storage at 20^{o}C. Plates were checked every day for the first week and once per week there after using a Leica stereozoom benchtop microscope.

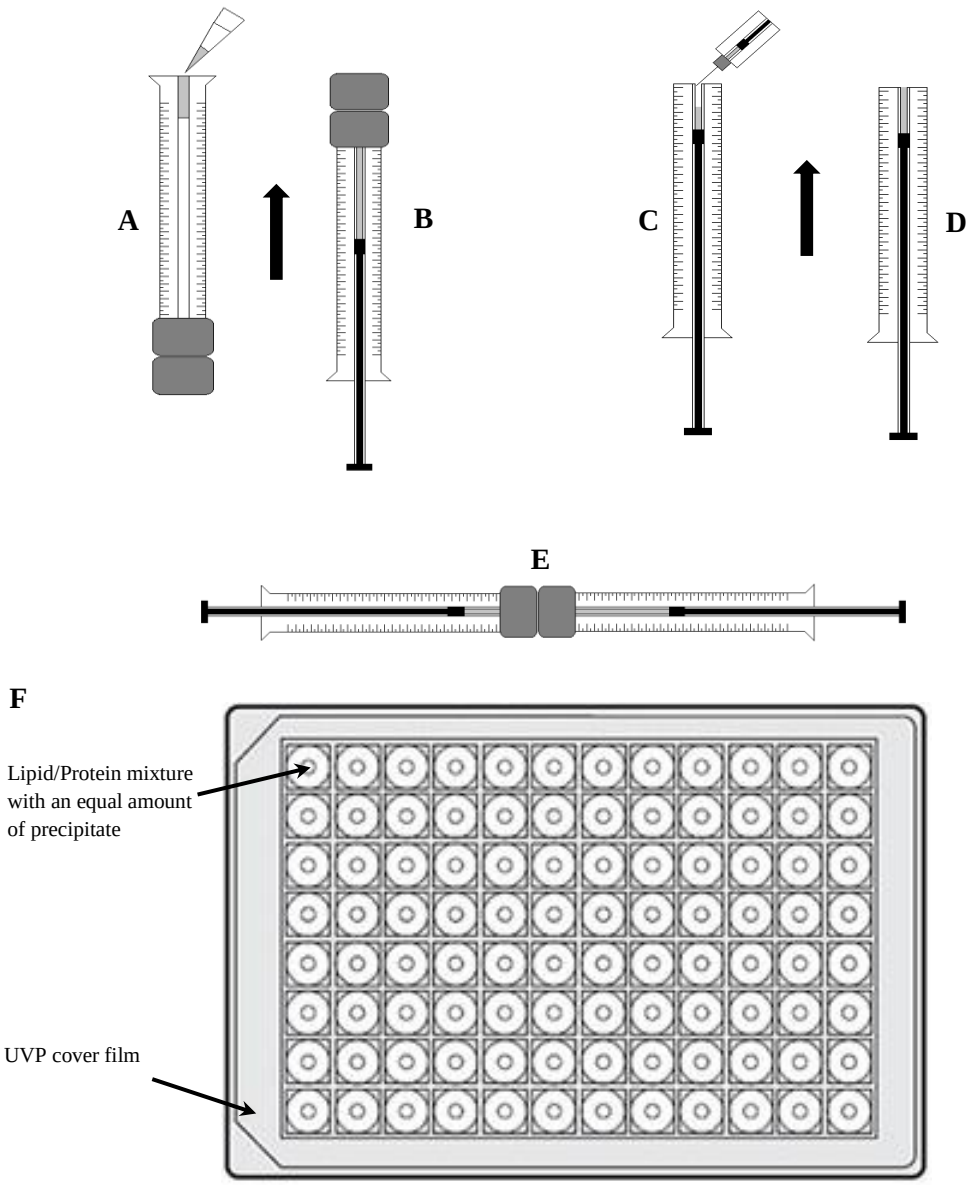


Figure 2.9. Protein preparation and set up of lipidic cubic phase crystallisation. A summary of the basic steps to reconstitute a protein sample into a lipid environment. **A.** The empty gas tight syringe was weighed and 30 mg of molten MO was added down the barrel. **B.** The plunger was replaced and the lipid pushed into the barrel. The loaded syringe and coupling unit was weighed to deduce the amount of MO used. **C**. To a second gas tight syringe the required amount of protein solution was added to the top of the syringe, in this case 20 µl at 10 mg/ml**. D.** The protein was moved up in the syringe in order to remove air bubbles. **E.** Both syringes were attached using the coupling unit and the contents passed between the two until a homogenous phase was achieved. **F**. Plate set up, flat 0.1mm deep 96 well plates were utilised. Each drop contained 0.5 µl of protein/lipid mix overlaid with 0.5 µl of any given precipitate. The plate was then sealed using UPV cover film and stored at 20ºC. Figure adapted from methodology produced by the Cherezov Lab, University of Southern California.

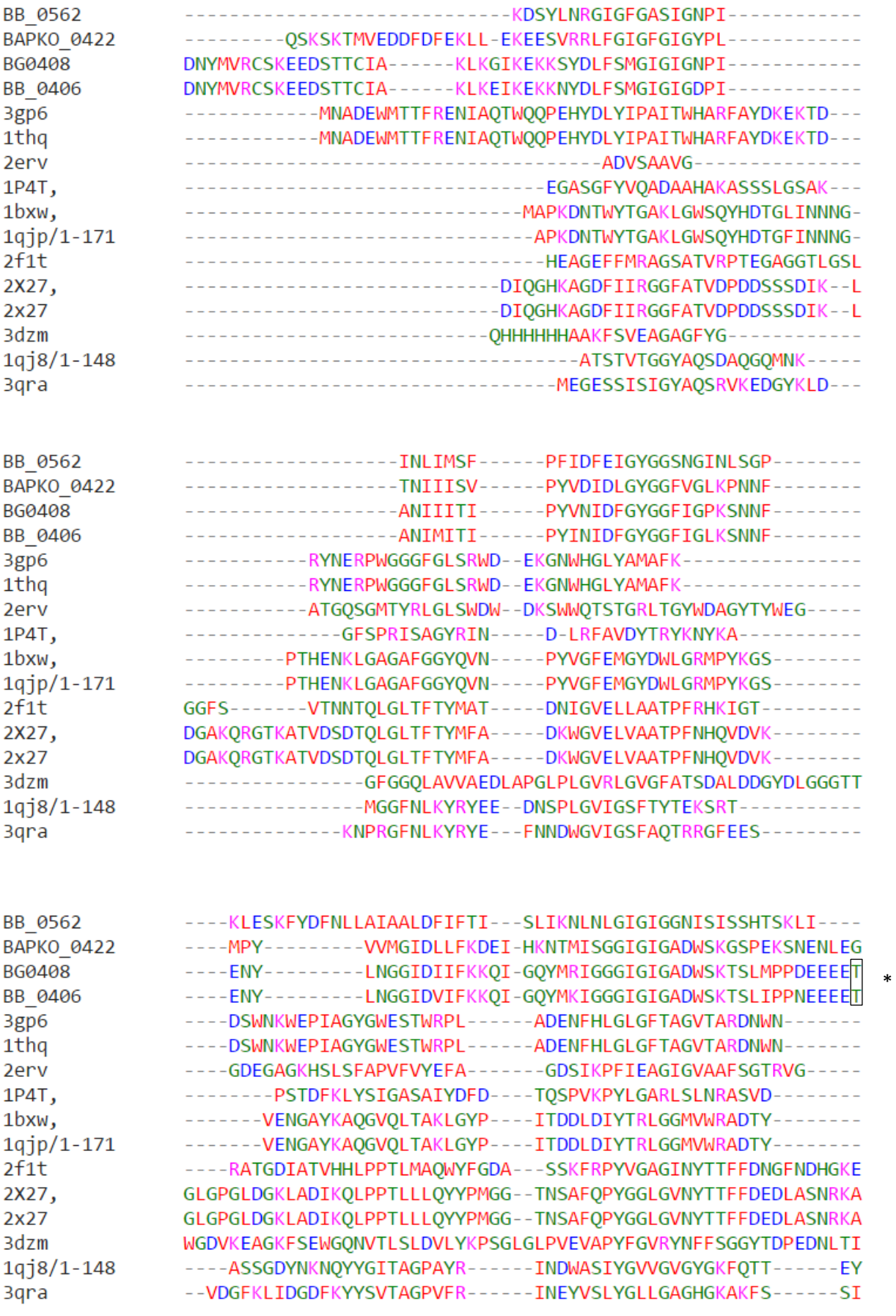
2.2.10 In silico characterisation and homology modelling of OMPs

Homology models for *Borrelia* OMPs were generated using Chimera (Pettersen et al., 2004) and Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000) alongside TMpred - ExPASy (Hofmann& Stoffel, 1993) in order to predict membrane spanning alignments. Although the modelling program PHYRE2 (Kelley, 2015) was the preferred method for homology modelling and was used for the *Borrelia* SodA project the program was unsuitable for OMP modelling. As the *Borrelia* OMPs held no significant level of sequence similarity to any known proteins templates were identified using a novel HMM approach (section 3.1). These templates shared very remote sequence similarity and a manual alignment of the templates was required which could not be achieved using PHYRE2 (Kelley, 2015). For this reason Chimera (Pettersen et al., 2004) and Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000) offered a powerful alternative.

A wide selection of 8-stranded beta barrel proteins with known structures (table 2.17) were aligned and used for predictions. The sequence alignment used for generation of homology models for *Borrelia* OMPs is shown in figure 2.10.

Protein	Organism	PDB code	Reference
PagP	*E. coli*	3GP6	Cuesta-Seijo et al., 2010
PagP	*E. coli*	1THQ	Ahn et al., 2004
PagL	*P. aeruginosa*	2ERV	Rutten et al., 2006
NspA	*N. meningitides*	1P4T	Vandeputte-Rutten et al., 2003
OmpA transmembrane domain	*E. coli*	1BXW	Pautsch and Schulz, 1998
OmpA transmembrane domain	*E. coli*	1QJP	Pautsch and Schulz, 2000
OmpW	*E. coli*	2F1T	Hong et al., 2006
Oprg	*P. aeruginosa*	2X27	Touw et al., 2010
Major outer membrane protein	*T. thermophilus*	3DZM	Brosig et al., 2009
OmpX	*E. coli*	1QJ8	Vogt and Schulz, 1999
Ail	*Y. pestis*	3QRA	Yamashita et al., 2011

Table 2.17. Bacterial OMPs used as templates for homology modelling. Protein sequences were manually aligned using Chimera (Pettersen et al., 2004) and Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000) and TMpred to provide a template for homology modelling.


BB_0562 ----------KDSYLNRGIGFGASIGNPI----------
BAPKO_0422 ---------QSKSKTMVEDDFDFEKLL-EKEESVRRLFGIGFGIGYPL----------
BG0408 DNYMVRCSKEEDSTTCIA------KLKGIKEKKSYDLFSMGIGIGNPI----------
BB_0406 DNYMVRCSKEEDSTTCIA------KLKEIKEKKNYDLFSMGIGIGDPI----------
3gp6 ----------MNADEWMTTFRENIAQTWQQPEHYDLYIPAITWHARFAYDKEKTD---
1thq ----------MNADEWMTTFRENIAQTWQQPEHYDLYIPAITWHARFAYDKEKTD---
2erv ----------ADVSAAVG----------
1P4T, ----------EGASGFYVQADAAHAKASSSLGSAK---
1bxw, ----------MAPKDNTWYTGAKLGWSQYHDTGLINNNG-
1qjp/1-171 ----------APKDNTWYTGAKLGWSQYHDTGFINNNG-
2f1t ----------HEAGEFFMRAGSATVRPTEGAGGTLGSL
2X27, ----------DIQGHKAGDFIIRGGFATVDPDDSSSDIK--L
2x27 ----------DIQGHKAGDFIIRGGFATVDPDDSSSDIK--L
3dzm ----------QHHHHHHAAKFSVEAGAGFYG----------
1qj8/1-148 ----------ATSTVTGGYAQSDAQGQMNK-----
3qra ----------MEGESSISIGYAQSRVKEDGYKLD---
BB_0562 ----------INLIMSF------PFIDFEIGYGGSNGINLSGP--------
BAPKO_0422 ----------TNIIISV------PYVDIDLGYGGFVGLKPNNF--------
BG0408 ----------ANIIITI------PYVNIDFGYGGFIGPKSNNF--------
BB_0406 ----------ANIMITI------PYINIDFGYGGFIGLKSNNF--------
3gp6 ----------RYNERPWGGGFGLSRWD--EKGNWHGLYAMAFK----------
1thq ----------RYNERPWGGGFGLSRWD--EKGNWHGLYAMAFK----------
2erv ----------ATGQSGMTYRLGLSWDW--DKSWWQTSTGRLTGYWDAGYTYWEG-----
1P4T, ----------GFSPRISAGYRIN-----D-LRFAVDYTRYKNYKA----------
1bxw, ---------PTHENKLGAGAFGGYQVN-----PYVGFEMGYDWLGRMPYKGS--------
1qjp/1-171 ---------PTHENKLGAGAFGGYQVN-----PYVGFEMGYDWLGRMPYKGS--------
2f1t GGFS-------VTNNTQLGLTFTYMAT-----DNIGVELLAATPFRHKIGT---------
2X27, DGAKQRGTKATVDSDTQLGLTFTYMFA-----DKWGVELVAATPFNHQVDVK--------
2x27 DGAKQRGTKATVDSDTQLGLTFTYMFA-----DKWGVELVAATPFNHQVDVK--------
3dzm ----------GFGGQLAVVAEDLAPGLPLGVRLGVGFATSDALDDGYDLGGGTT
1qj8/1-148 ----------MGGFNLKYRYEE--DNSPLGVIGSFTYTEKSRT----------
3qra ----------KNPRGFNLKYRYE---FNNDWGVIGSFAQTRRGFEES---------
BB_0562 ----KLESKFYDFNLLAIAALDFIFTI---SLIKNLNLGIGIGGNISISSHTSKLI----
BAPKO_0422 ----MPY---------VVMGIDLLFKDEI-HKNTMISGGIGIGADWSKGSPEKSNENLEG
BG0408 ----ENY---------LNGGIDIIFKKQI-GQYMRIGGGIGIGADWSKTSLMPPDEEEET *
BB_0406 ----ENY---------LNGGIDVIFKKQI-GQYMKIGGGIGIGADWSKTSLIPPNEEEET
3gp6 ----DSWNKWEPIAGYGWESTWRPL------ADENFHLGLGFTAGVTARDNWN-------
1thq ----DSWNKWEPIAGYGWESTWRPL------ADENFHLGLGFTAGVTARDNWN-------
2erv ----GDEGAGKHSLSFAPVFVYEFA-------GDSIKPFIEAGIGVAAFSGTRVG-----
1P4T, --------PSTDFKLYSIGASAIYDFD----TQSPVKPYLGARLSLNRASVD--------
1bxw, -------VENGAYKAQGVQLTAKLGYP----ITDDLDIYTRLGGMVWRADTY--------
1qjp/1-171 -------VENGAYKAQGVQLTAKLGYP----ITDDLDIYTRLGGMVWRADTY--------
2f1t ----RATGDIATVHHLPPTLMAQWYFGDA---SSKFRPYVGAGINYTTFFDNGFNDHGKE
2X27, GLGPGLDGKLADIKQLPPTLLLQYYPMGG--TNSAFQPYGGLGVNYTTFFDEDLASNRKA
2x27 GLGPGLDGKLADIKQLPPTLLLQYYPMGG--TNSAFQPYGGLGVNYTTFFDEDLASNRKA
3dzm WGDVKEAGKFSEWGQNVTLSLDVLYKPSGLGLPVEVAPYFGVRYNFFSGGYTDPEDNLTI
1qj8/1-148 ----ASSGDYNKNQYYGITAGPAYR------INDWASIYGVVGVGYGKFQTT------EY
3qra --VDGFKLIDGDFKYYSVTAGPVFR------INEYVSLYGLLGAGHGKAKFS------SI

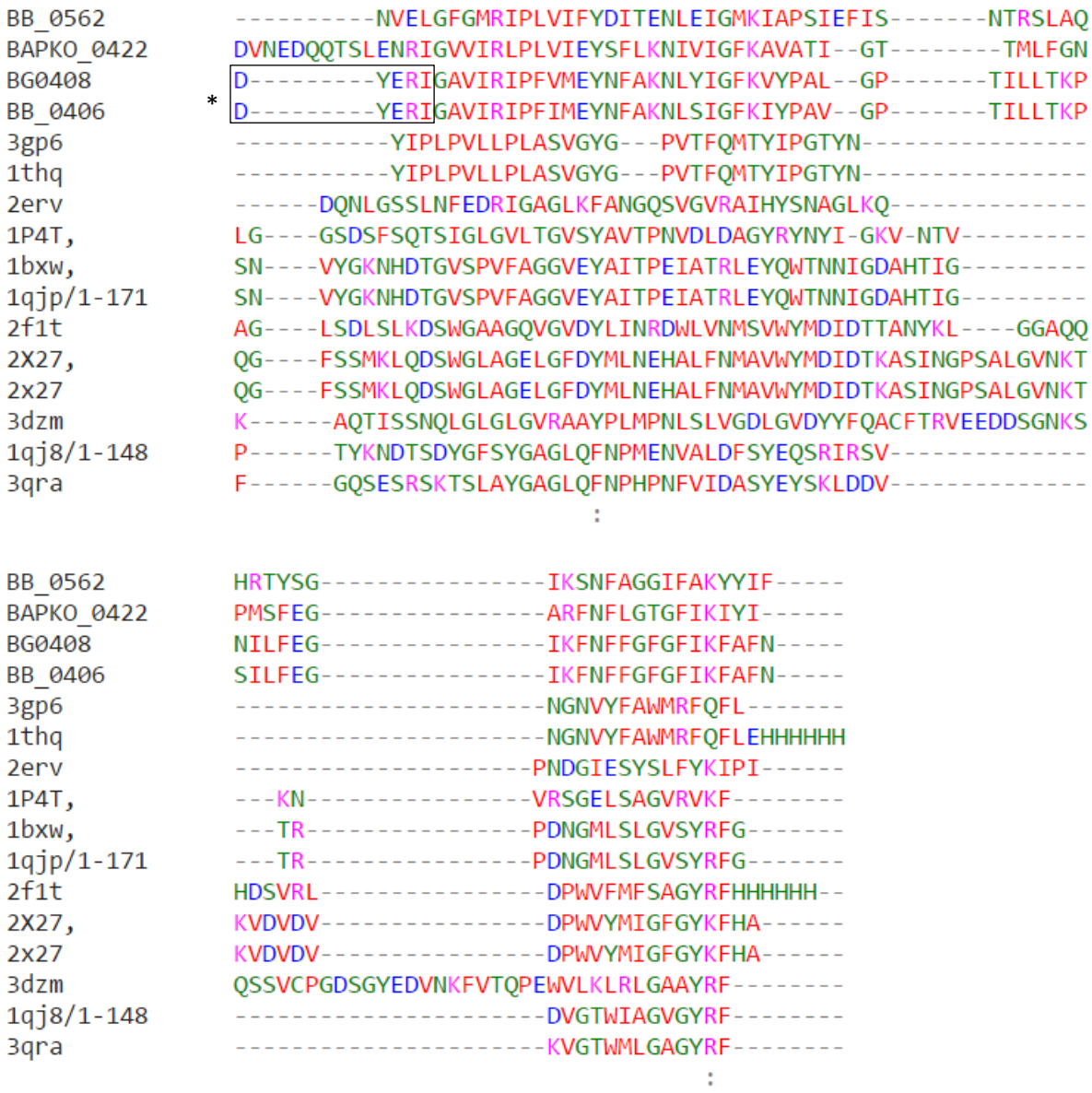


Figure 2.10. Sequence alignment for construction of homology models for *Borrelia* OMPs. The templates and alignment used for homology modelling of *Borrelia* OMPs using Chimera (Pettersen et al., 2004) and Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000). Pink residues are positive, blue are negative, red are non-polar and green any other. * Indicates a potential lipocalin domain TDY—I within BB_0406 and BG0408.

Chapter 3: Results - *Borrelia* OMPA-like proteins

3.1 Identification of possible OmpA-like proteins in *B. burgdorferi* s.l

Previous work undertaken by Dr Adam Dyer identified a number of possible OmpA-like proteins within *B. burgdorferi s.l.*(Thesis, 2013, University of Huddersfield). As traditional protein BLAST (Altschul et al., 1990) searches of known *Borrelia* genomes using *E. coli* OmpA as a search term proved insensitive to detect homologues based on sequence similarity a Hidden Markov Model system was employed. The approach involved the selection of known OmpA-like proteins across a broad range of bacterial species and their sequences aligned. Generation of the multiple sequence alignments allowed the identification of the most conserved residues and also the regions that contain a higher rate of mutation, insertion or deletion. By recognising shared key residues within a protein family and building a HMM profile of multiple sequence alignments, the sensitivity of discovering more remote homologues was increased. From the large number of possible candidates identified any protein with a putative known function was removed alongside anything of a molecular weight much larger or smaller than that of known OmpA-like proteins. Following this exclusion a small number of proteins across *B. burgdorferi*, *B. afzelii* and *B. garinii* were identified as possible OmpA-like proteins and are summarised in table 1.10 within section 1.7.1.

3.2 Amplification and cloning of *B. burgdorferi* s.l OMPA-like genes.

3.2.1 Designing primers and signal sequence analysis

Four primers were designed for each gene target, two for initial amplification and a second pair for the addition of BamHI and NotI restriction sites. It was decided that the signal sequence of each protein was to be removed in order to prevent translocation of recombinant proteins into the *E. coli* membrane and to force the formation of inclusion bodies. In order to achieve this, the three target sequences were sent to Signal P, their signal sequences were deduced with data shown in figure 3.1 and amplification primers were designed to bind after this sequence.

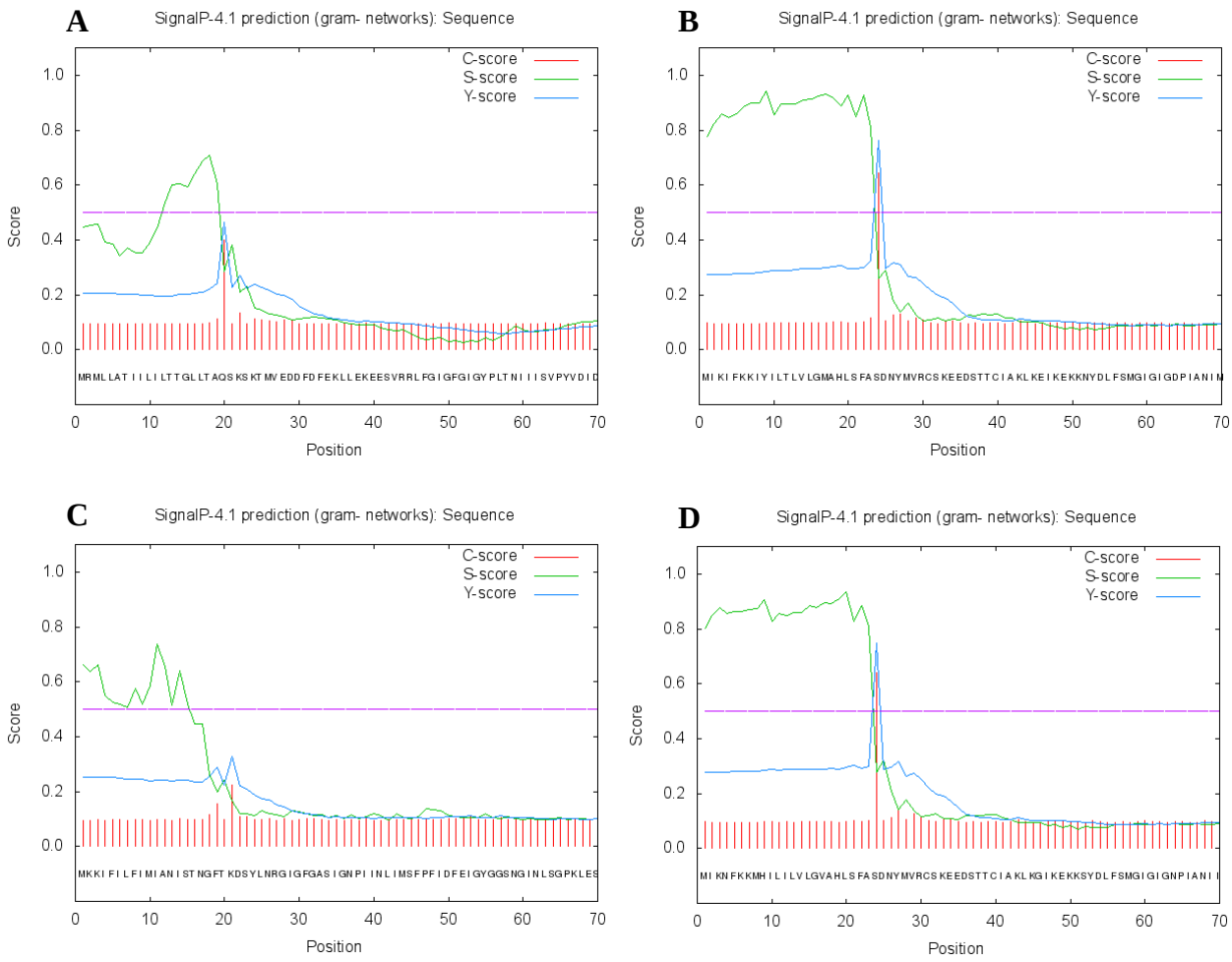


Figure 3.1. Results of SignalP analysis for Omp-A like proteins. All images were generated by SignalP 4.1 (Petersen et al., 2011). The beginning of each proteins amino acid runs across the bottom with the red lines indicating a c-score which is used to assess the likelihood of a cleavage site at each amino acid position. The green or s-score, represents the probability of each residue being part of the single sequence and the blue line is the y-score or a combination of the other two values in order to predict the site more accurately than using the c score alone. **A)** BAPKO_0422. – Signal sequence does not meet the threshold value set by Signal P for Gram-negative proteins. **B)** BB_0406 – Signal sequence present and exceeds the threshold value. **C)** BB_0562 – No signal sequence identified. **D)** BG0408 – Signal sequence present and exceeds the threshold value.

3.2.2 Amplification of the target genes

Following identification of possible OmpA-like genes in *B. afzelii*, *B. burgdorferi* and *B. garinii*, targets were amplified using the primers described in tables 2.1 and 2.2. As DNA kindly supplied by Dr Volker Fingerle (German National Reference Centre for *Borrelia*) was of high quality for each *Borrelia* strain, nested PCR was not required and a standard method of PCR was employed.

Initial denaturation		The reaction mix was heated to 95ºC for 5 minutes.
Amplification cycles	**Denaturation**	The reaction mix was heated to 95ºC for 30 seconds.
	Annealing	The reaction temperature was lowered for 30 seconds BB_0562 – 52ºC, BB_0406 – 52ºC and BG0408 – 51ºC.
	Extension	The reaction mix was heated to 72ºC for 1 minute.
Final extension		The reaction mix was heated to 72ºC for 5 minutes

Table 3.1. Experimentally determined PCR parameters. The PCR reactions contain three distinct stages including an initial denaturation, the amplification cycles and a final extension stage. The temperatures for each stage were experimentally determined to optimise the final product yield.

Initial amplification was carried out at 4 and 5 ºC below than the lowest melting temperature for each primer pair. This was deemed to be 49 ºC and 50 ºC for BB_0562 and 48 ºC and 49ºC for both BB_0406 and BG_0408. Unfortunately this failed to yield product and the annealing temperature was increased for all reactions to 51 and 52 ºC (table 3.1). Product was obtained for all three targets following the increase of the annealing temperature with results shown in figure 3.3. It became apparent that the reactions were yielding two products of a similar size illustrated by the two bands shown on the agarose gel in figure 3.3, these products were both close to the expected size of the target genes at around 600-700bp. The 52 ºC PCR reaction for BB_0562 and BB_0406 alongside the 51 ºC reaction for BG_0408 were purified and subjected to restriction PCR. Restriction sites were added to all three genes successfully with a further round of PCR using 35 cycles and the appropriate primer combination.

The restriction primers were designed to produce a product as similar as possible to the native gene minus the deduced signal sequence. In order to aid restriction digests non-native base pairs were added during restriction PCR. These six additional bases (AATTAT) were added to

the 5' end of the primer and were followed by the required restriction site and then a complementary sequence for hybridisation to the amplified gene product as summarised in figure 3.2 In some cases a single base (G) was added following the restriction site in order to maintain the reading frame however the codon was always checked to ensure a stop codon was not misplaced. The reverse primer also contained the complement to the native gene stop codon.

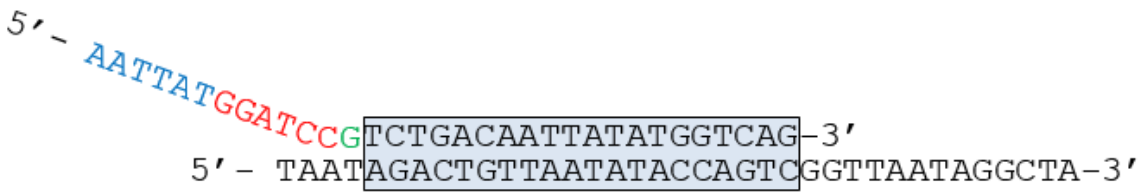


Figure 3.2. Restriction PCR - An example of a forward overhang primer. The lower strand is the initial PCR product from the first round of amplification, the upper strand is the restriction PCR primer. The base pairs within the grey box represent the portion of the primer which is complementary to the PRC product and hybridises during the annealing stage. The remaining 5' end of the primer outside if the grey box is non-complementary and overhangs the DNA template. This overhand consists of a single base (green) added in order to keep the forward reading frame, the restriction site for in this case BamHI (red) and finally six additional bases (blue) to assist subsequent restriction digests during the cloning process.

Interestingly the restriction PCR provided a single product unlike the original amplification which gave two as previously discussed and shown in figure 3.4. It is likely that the initial amplification gave the same product although one was extended further past the gene possibly due non-specific binding of the primer during the initial amplification explaining the difference in length. This was later rectified during the PCR addition of the restriction sites. During this reaction only the product of the correct length could be used as a template for amplification, as the longer product would not have been a complimentary match to the designed primers.

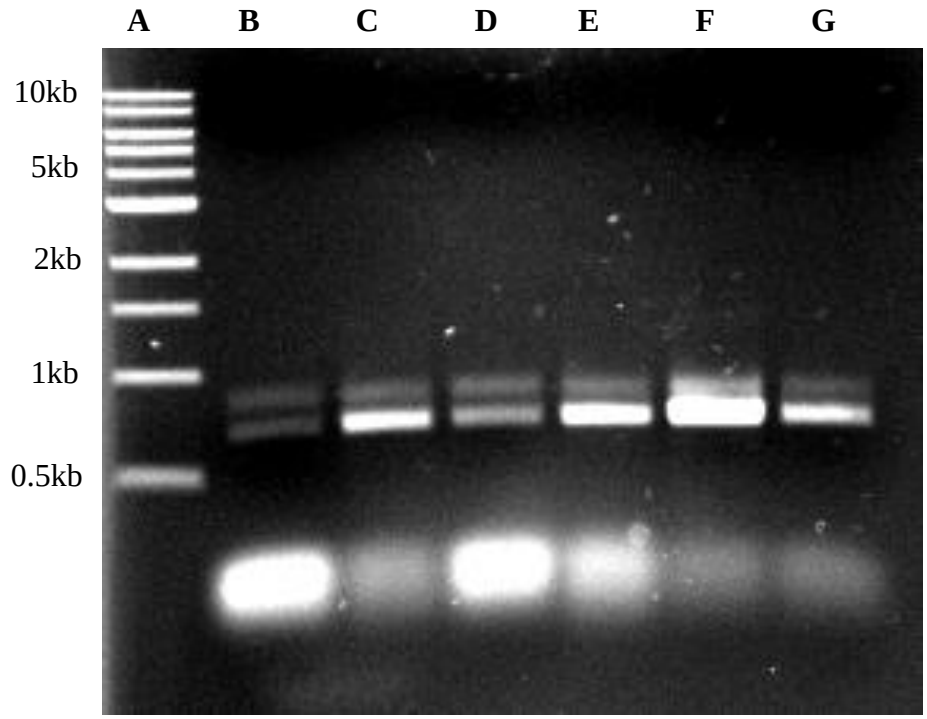


Figure 3.3. Agarose electrophoresis of the initial amplification of OmpA-like proteins. PCR products generated from the primers listed in table 2.1 and 2.2 and ran on a 1% w/v agarose gel in 1X TAE buffer (table 2.5). A) NEB 1kb ladder. B) BB_0562 at 51ºC. C) BB_0562 at 52 ºC. D) BB_0406 at 51 ºC. E) BB_0406 at 52 ºC. F) BG_0406 at 51 ºC. G) BG_0408 at 52ºC.

The annealing temperatures selected for restriction PCR covered only two temperatures for all three targets as product was attained upon the first attempt. Results from a gel electrophoresis for the restriction PCR are shown in figure 3.4 and show that strong amplification was achieved at 52ºC for BB_0562 and BB_0406 and at 51ºC for BG0408.

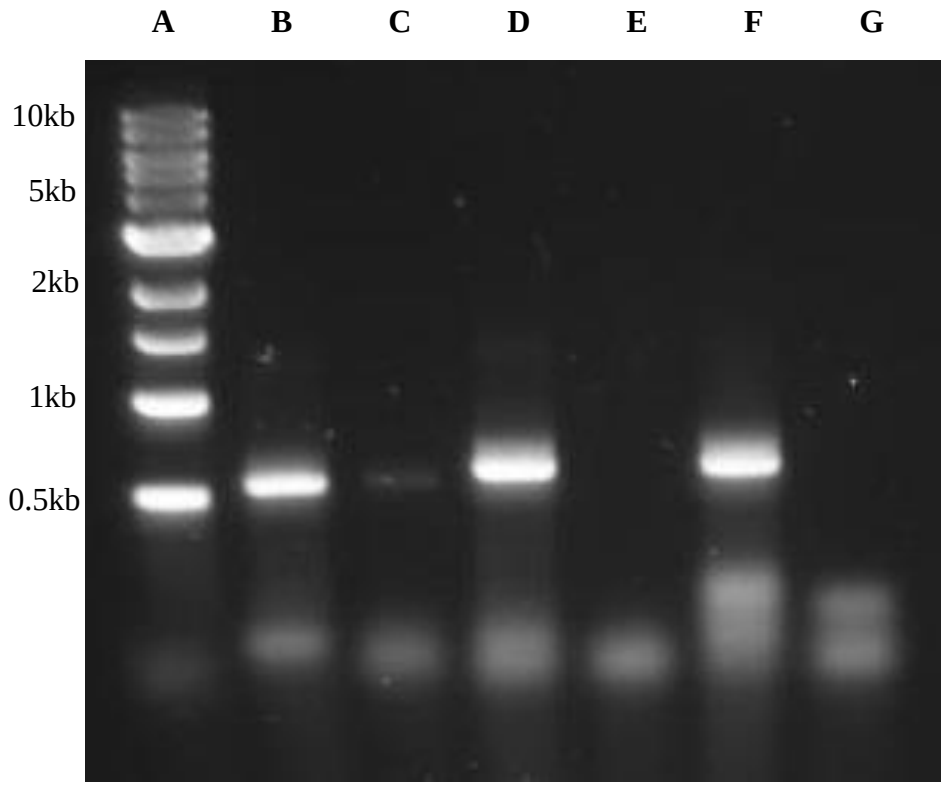


Figure 3.4. Agarose electrophoresis of the restriction PCR of OmpA-like proteins. PCR products generated from the primers discussed in tables 2.1 and 2.2 and ran on a 1% w/v agarose gel in 1X TAE buffer (table 2.5). A) NEB 1kb ladder. B) BB_0562 at 51ºC. C) BB_0562 52ºC. D) BB_0406 at 51ºC. E) BB_0406 52ºC. F) BG_0406 at 51 ºC. G) BG_0408 at 51ºC.

3.2.3 Generation of gene constructs

Following addition of the required restriction sites, the gene inserts and the pET47 vector were sequentially digested with appropriate restriction enzymes in order to liberate complimentary restriction sites for ligation (section 2.2.2). Following initial problems with inefficient digestion all digests were carried out overnight and all ligation reactions subjected to a final digest with the restriction enzyme SalI. During initial cloning experiments inefficient digestion of PCR products and the vector became problematic and led to the formation of false positive colonies following transformation. These false positives were caused by either uncut vector or vector cut with only one restriction enzyme which became re-circularised during ligation. These 'empty' vectors did not contain the desired target genes yet were passing through the experimental procedure and obscuring the selection of true transformants. This issue was rectified by implementing a final restriction digest, where all ligation reactions were digested with SalI following the heat inactivation of T4 ligase. The SalI site was not present in any of

the target gene inserts and the site is lost following the full digestion of pET47 with BamHI and NotI, making the site a desirable target for the destruction of any 'empty' vectors (section 2.2.2).

All ligation ratios (1:3, 1:5 and 1:10) appeared to provide positive clones although ligation reactions with a smaller amount of plasmid DNA (down to 5ng) were seen to be most successful at incorporating gene inserts into the pET47 vector most likely due to a higher efficiency of restriction due to the smaller amount of vector to digest (data not shown). Following the adaption of the protocol (section 2.2.2.4) to remove any re-circularised plasmid multiple clones were selected across all ligation ratios and were used to inoculate individual vials of 5 ml of LB supplemented with Kanamycin (50µg/ml) and incubated overnight at 37^{o}C for 12 hours with vigorous shaking. Plasmids were extracted from each culture using a QIAprep Spin Miniprep kit (Qiagen, UK) and used for restriction analysis.

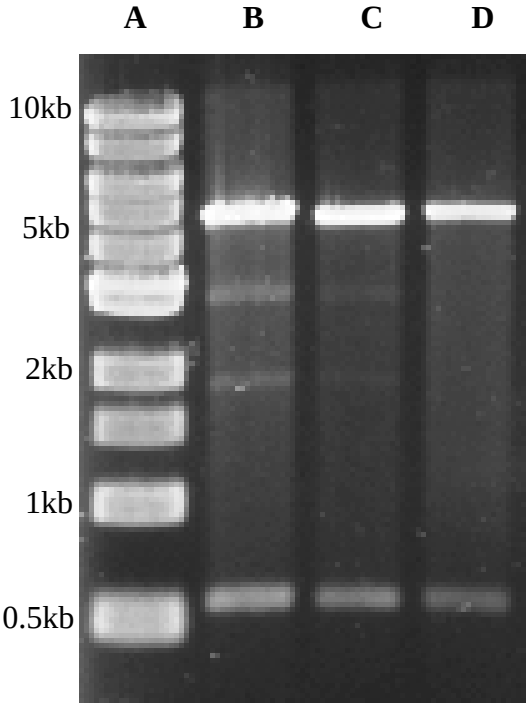


Figure 3.5. Agarose electrophoresis of the restriction analysis of OmpA-like constructs. Restriction analysis following an extended double digest with BamHI and NotI and ran on a 1% w/v agarose gel in 1X TAE buffer (table 2.5). A) NEB 1kb DNA ladder. B) Digested pET47-BB_0406. C) Digested pET47-BG_0408. D) Digested pET47- BB_0562.

Restriction analysis (section 2.2.2.6) was performed to identify the presence of a correct sized gene insert prior to sequencing. Constructs were digested with BamHI and NotI as previously described and the fragments analysed by gel electrophoresis. The gel analysis showed two bands for each construct following digestion, the first in the region of 5kb and the second at around ~600bp, gels are shown in figure 3.5. The analysis also showed two other bands for the BB_0406 construct. These additional bands are likely to be the result of star activity from BamHI during the extended digest in the NotI digest buffer as samples were not subjected to miniprep prior to the second digestion.

3.2.4 Sequencing of the generated constructs

Following confirmation that the constructs contained an insert at the expected size all vectors were sent to Source Bioscience (Rochdale, UK) for Sanger sequencing using T7 forward and reverse stock primers; sequences are shown in figure 3.6 with full sequencing data included in the appendix 1. Returned sequences were translated using EXPASY and the reading frame checked. Following translation the resulting peptide was run through Protein BLAST (Altschul et al., 1990) to ensure a match to the original gene. Sequence alignments for all three constructs can be seen in figure 3.7. It was noticed there were polymorphisms within the BB_0406 and BG0408 compared to the sequence from the NCBI database (https://www.ncbi.nlm.nih.gov/). Although these could have occurred during the cloning process it is most likely these polymorphisms are naturally found within the DNA provided for use as a template.

BB_0562

Nucleotide sequence

TATGGCACATCACCACCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCAGGATCCGAA
AGATTCATATTTAAATAGAGGAATTGGCTTTGGAGCAAGCATTGGAAATCCAATTATTAACTTAATAATGTCATT
TCCTTTCATTGATTTTGAAATTGGCTATGGTGGTAGTAATGGAATAAATCTATCAGGCCCCAAACTTGAATCAAA
ATTTTATGATTTTAATTTATTAGCAATAGCAGCACTTGATTTCATTTTTACAATATCTTTGATAAAAAATTTAAA
TTTAGGAATTGGAATAGGAGGAAATATAAGCATATCGTCTCACACATCTAAATTAATAAATGTAGAATTAGGATT
TGGAATGAGAATTCCATTGGTTATTTTTTACGACATTACAGAAAATTTAGAAATAGGTATGAAAATAGCACCTTC
AATAGAATTCATCTCAAATACAAGGTCTCTTGCTCAACATAGAACCTATTCGGGCATAAAATCAAACTTTGCTGG
GGGAATATTTGCTAAGTACTATATCTTTTAACA

Amino acid sequence

MAHHHHHHSAALEVLFQGPGYQDPKDSYLNRGIGFGASIGNPIINLIMSFPFIDFEIGYGGSNGINLSGPKLESK
FYDFNLLAIAALDFIFTISLIKNLNLGIGIGGNISISSHTSKLINVELGFGMRIPLVIFYDITENLEIGMKIAPS
IEFISNTRSLAQHRTYSGIKSNFAGGIFAKYYIF

BB_0406

Nucleotide sequence

TATGGCACATCACCACCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCAGGATCCGTC
TGACAATTATATGGTCAGATGCAGCAAGGAAGAAGATTCAACCACCTGTATCGCAAAGCTTAAAGAAATAAAAGA
AAAGAAAGATTATGACTTATTTTCAATGGGCATTGGAATAGGAGATCCTATTGCAAATATTATGATTACAATTCC
TTATATAAATATTGATTTTGGATGTGGAGGTTTTATTGGCCTTAAGTCAAACAATTTTGAAAATTATCTAAATGG
TGGAATAGACGTTATTTTTAAAAAGCAAATTGGACAATATATGAAAATTGGCGGCGGCATTGGAATAGGTGCGGA
TTGGTCAAAAACATCCCTTATACCCCCTAATGAAGAAGAAGAAACTGATTATGAGAGAATAGGCGCTGTTATAAG
AATTCCTTTTATAATGGAATATAATTTTGCAAAAAATTTATCCATAGGATTCAAAATTTATCCTGCAGTAGGGCC
AACAATATTACTAACAAAACCCAGCATTTTATTTGAAGGAATTAAATTCAATTTTTTTGGATTTGGATTCATAAA
ATTTGCATTTAATTA

Amino acid sequence

MAHHHHHHSAALEVLFQGPGYQDPSDNYMVRCSKEEDSTTCIAKLKEIKEKKDYDLFSMGIGIGDPIANIMITIP
YINIDFGCGGFIGLKSNNFENYLNGGIDVIFKKQIGQYMKIGGGIGIGADWSKTSLIPPNEEEETDYERIGAVIR
IPFIMEYNFAKNLSIGFKIYPAVGPTILLTKPSILFEGIKFNFFGFGFIKFAFN

BG_0408

Nucleotide sequence

TATGGCACATCACCACCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCAGGATCCGTC
TGACAATTATATGGTCAGACGCAGCAAGGAAGAAGATTCAACAACCTGTATCGCAAAGCTTAAAGGCATAAAAGA
AAAAAAAAGTTATGACTTATTTTCAATGGGTATTGGAATAGGCAATCCTATTGCAAACATAATAATTACAATTCC
TTATATAAATATTGATTTTGGATATGGGGGTTTTATTGGTCCTAAGTCAAATAATTTTGAAAATTATCTAAACGG
CGGAATAGATATTGTTTTTAAAAAACAAATTGGACAATACATGAGAATCGGTGGCGGCATTGGAATAGGTGCAGA
TTGGTCAAAAACATCCCTCATACCTCCTGACGAAGAGAAAGAGACTGATTATGAGAGAATAGGCGCTGTTATAAG
AATTCCTTTTGTAATGGAATATAACTTTGCAAAAAATTTGTACATAGGATTCAAACTTTACCCTGCACTAGGGCC
AACAATATTGCTAACAAAGCCAAAAATTTTATTTGAAGGAATTAAATTCAATTTTTTTGGAT

Amino acid sequence

*M*AHHHHHHSAALEVLFQGPGYQDPSDNYMVRRSKEEDSTTCIAKLKGIKEKKSYDLFSMGIGIGNPIANIIITIP
YINIDFGYGGFIGPKSNNFENYLNGGIDIVFKKQIGQYMRIGGGIGIGADWSKTSLIPPDEEKETDYERIGAVIR
IPFVMEYNFAKNLYIGFKLYPALGPTILLTKPKILFEGIKFNFFGFGFIKFAFN

Figure 3.6. Sequence data of OMP-pET47 constructs and resulting translations. DNA sequences and EXPASY translations following Sanger sequencing of the OMP-pET47 constructs using T7 promoter and terminator primers. Red shows the start codon, green the his-tag region, orange the HRV3C cleavage site and the gene of interest is shown in dark blue.

BB_0562

```
  1 --------------------KDSYLNRGIGFGASIGNPIINLIMSFPFID     30
                       |||||||||||||||||||||||||||||||
  1 MKKIFILFIMIANISTNGFTKDSYLNRGIGFGASIGNPIINLIMSFPFID     50

 31 FEIGYGGSNGINLSGPKLESKFYDFNLLAIAALDFIFTISLIKNLNLGIG     80
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
 51 FEIGYGGSNGINLSGPKLESKFYDFNLLAIAALDFIFTISLIKNLNLGIG    100

 81 IGGNISISSHTSKLINVELGFGMRIPLVIFYDITENLEIGMKIAPSIEFI    130
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
101 IGGNISISSHTSKLINVELGFGMRIPLVIFYDITENLEIGMKIAPSIEFI    150

131 SNTRSLAQHRTYSGIKSNFAGGIFAKYYIF    160
    ||||||||||||||||||||||||||||||
151 SNTRSLAQHRTYSGIKSNFAGGIFAKYYIF    180
```

BB_0406

```
  1 -----------------------SDNYMVRCSKEEDSTTCIAKLKEIKEK     27
                            ||||||||||||||||||||||||||||
  1 MIKIFKKIYILTLVLGMAHLSFASDNYMVRCSKEEDSTTCIAKLKEIKEK     50

 28 KDYDLFSMGIGIGDPIANIMITIPYINIDFGCGGFIGLKSNNFENYLNGG     77
    |:|||||||||||||||||||||||||||||.||||||||||||||||||||
 51 KNYDLFSMGIGIGDPIANIMITIPYINIDFGYGGFIGLKSNNFENYLNGG    100

 78 IDVIFKKQIGQYMKIGGGIGIGADWSKTSLIPPNEEEETDYERIGAVIRI    127
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
101 IDVIFKKQIGQYMKIGGGIGIGADWSKTSLIPPNEEEETDYERIGAVIRI    150

128 PFIMEYNFAKNLSIGFKIYPAVGPTILLTKPSILFEGIKFNFFGFGFIKF    177
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
151 PFIMEYNFAKNLSIGFKIYPAVGPTILLTKPSILFEGIKFNFFGFGFIKF    200

178 AFN    180
    |||
201 AFN    203
```

BG0408

```
  1 -----------------------SDNYMVRRSKEEDSTTCIAKLKGIKEK     27
                            |||||||.||||||||||||||||||||
  1 MIKNFKKMHILILVLGVAHLSFASDNYMVRCSKEEDSTTCIAKLKGIKEK     50

 28 KSYDLFSMGIGIGNPIANIIITIPYINIDFGYGGFIGPKSNNFENYLNGG     77
    |||||||||||||||||||||||||:||||||||||||||||||||||||||
 51 KSYDLFSMGIGIGNPIANIIITIPYVNIDFGYGGFIGPKSNNFENYLNGG    100

 78 IDIVFKKQIGQYMRIGGGIGIGADWSKTSLIPPDEEKETDYERIGAVIRI    127
    |||:|||||||||||||||||||||||||:|||||:||||||||||||||
101 IDIIFKKQIGQYMRIGGGIGIGADWSKTSLMPPDEEEETDYERIGAVIRI    150

128 PFVMEYNFAKNLYIGFKLYPALGPTILLTKPKILFEGIKFNFFGFGFIKF    177
    |||||||||||||||||:||||||||||||.||||||||||||||||||||
151 PFVMEYNFAKNLYIGFKVYPALGPTILLTKPNILFEGIKFNFFGFGFIKF    200

178 AFN    180
    |||
201 AFN    203
```

Figure 3.7. Protein BLAST results of translated OMP sequence data. Following translations of the OMP-pET47 constructs the gene inserts were translated using EXPASY and used for protein BLAST (Altschul et al., 1990) searching. The top sequence is the query or the gene insert with the actual NCBI (https://www.ncbi.nlm.nih.gov/) result underneath. Note that the query sequences are shorter due to the removal of the signal sequence.

3.3 Recombinant expression of BAPKO_0422, BB_0562, BB_0406 and BG0408

3.3.1 Codon bias

All four gene sequences (BAPKO_0422, BB_0562, BB_0406 and BG_0408) were analysed for rare codons using the Rare Coon Calculator (RaCC) from NIH MBI Laboratory for Structural Genomics and Proteomics. Unsurprisingly all of the genes contained a number of codons not frequently used with the expression system *E. coli* K12 and is further discussed in table 3.2. In order to achieve a high level of expression without premature termination, Rosetta (DE3) *E. coli* cells were selected for expression. The expression cells contained additional tRNAs for AGG, AGA, AUA, CUA, CCC, CGA, all of which were required for translation of the *Borrelia* proteins. These extra tRNAs were encoded on a chloramphenicol resistant plasmid and the strain gave high levels of overexpression.

3.3.2 Autoinduction vs IPTG control

Initial trials of autoinduction (section 2.2.3.2) failed to produce high levels of recombinant protein following 16 hours of incubation at 37^{o}C at 200rpm when compared to induction with IPTG. This was rectified by increasing the incubation time to 20 hours. Protein yield was significantly higher following the extended incubation and much higher than expression under IPTG control likely due to the higher culture density achieved during autoinduction (Grabski et al., 2005).

			Rare Arg codons AGG, AGA, CGA	Rare Leu codons CTA	Rare Ile codons ATA	Rare Pro codons CCC
B. burgdorferi B31	BB_0562 Predicted outer membrane β-barrel	gtg aaa aaa att ttt ATA ttg ttt atc atg att gca aac ATA tct aca aat ggt ttt aca aaa gat tca tat tta aat AGA gga att ggc ttt gga gca agc att gga aat cca att att aac tta ATA atg tca ttt cct ttc att gat ttt gaa att ggc tat ggt ggt agt aat gga ATA aat CTA tca ggc CCC aaa ctt gaa tca aaa ttt tat gat ttt aat tta tta gca ATA gca gca ctt gat ttc att ttt aca ATA tct ttg ATA aaa aat tta aat tta gga att gga ATA gga gga aat ATA agc ATA tcg tct cac aca tct aaa tta ATA aat gta gaa tta gga ttt gga atg AGA att cca ttg gtt att ttt tac gac att aca gaa aat tta gaa ATA ggt atg aaa ATA gca cct tca ATA gaa ttc atc tca aat aca AGG tct ctt gct caa cat AGA acc tat tcg ggc ATA aaa tca aac ttt gct ggg gga ATA ttt gct aag tac tat atc ttt taa	4	1	16	1
B. burgdorferi B31	BB_0406 Predicted outer membrane β-barrel	atg ATA aaa att ttt aaa aaa ATA tac att tta aca tta gta tta ggt atg gca cac ctt tct ttt gca tct gac aat tat atg gtc AGA tgc agc aag gaa gaa gat tca acc acc tgt atc gca aag ctt aaa gaa ATA aaa gaa aag aaa aat tat gac tta ttt tca atg ggc att gga ATA gga gat cct att gca aat att atg att aca att cct tat ATA aat att gat ttt gga tat gga ggt ttt att ggc ctt aag tca aac aat ttt gaa aat tat CTA aat ggt gga ATA gac gtt att ttt aaa aag caa att gga caa tat atg aaa att ggc ggc ggc att gga ATA ggt gcg gat tgg tca aaa aca tcc ctt ATA CCC cct aat gaa gaa gaa gaa act gat tat gag AGA ATA ggc gct gtt ATA AGA att cct ttt ATA atg gaa tat aat ttt gca aaa aat tta tcc ATA gga ttc aaa att tat cct gca gta ggg cca aca ATA tta CTA aca aaa cca agc att tta ttt gaa gga att aaa ttc aat ttt ttt gga ttt gga ttc ATA aaa ttt gca ttt aat taa	3	2	14	1
B. garinii PBi	BG_0408 Predicted outer membrane β-barrel	atg ATA aaa aat ttt aaa aaa atg cat att tta ATA tta gta tta ggt gta gca cac ctt tct ttt gca tct gac aat tat atg gtc AGA tgc agc aag gaa gaa gat tca acc acc tgt atc gca aag ctt aaa ggc ATA aaa gaa aaa aaa agt tat gac tta ttt tca atg ggt att gga ATA ggt aat cct att gca aac atc ATA att aca att cct tat gta aat att gat ttt gga tat ggg ggt ttt att ggc cct aag tca aat aat ttt gaa aat tat CTA aac ggc gga ATA gat att att ttt aaa aaa caa att gga caa tac atg AGA atc ggt ggc ggt att gga ATA ggt gca gat tgg tca aaa aca tcc ctt atg cct cct gac gaa gag gaa gag act gat tat gag AGA ATA ggc gct gtt ATA AGA att cct ttt gta atg gaa tat aac ttt gca aaa aat tta tac ATA gga ttc aaa gtt tac cct gca CTA ggg cca aca ATA ttg CTA aca aag cca aac att tta ttt gaa gga att aaa ttc aat ttt ttt gga ttt gga ttc ATA aaa ttt gca ttt aat taa	4	3	16	2

Table 3.2. Rare codon analysis of *Borrelia* burgdorferi s.l. OmpA-like proteins. Prior to transformation of OMP-pET47 vectors for expression codon usage was calculated in order to select an appropriate expression strain of *E. coli*. Rare codons were identified using the Rare Coon Calculator (RaCC) from NIH MBI Laboratory for Structural Genomics and Proteomics

3.3.3 Western blot confirmation of expression

Following expression western blot analysis was used to confirm the presence of the recombinant protein by the means of immunoblotting for the polyhistidine tag (section 2.2.3.5). Western analysis demonstrated that expression of all four proteins was achieved and these recombinant proteins were of an expected size (20-23 kDa) as shown in figure 3.8. Prior to blotting protein samples were separated by reducing SDS-PAGE in order to reduce any oligomers however in some cases BB_0406 and BG_0408 appeared on blots with three/four bands possibly representing forms of dimer tetramer and octamers (data not shown) or oligomerisation from incomplete folding.

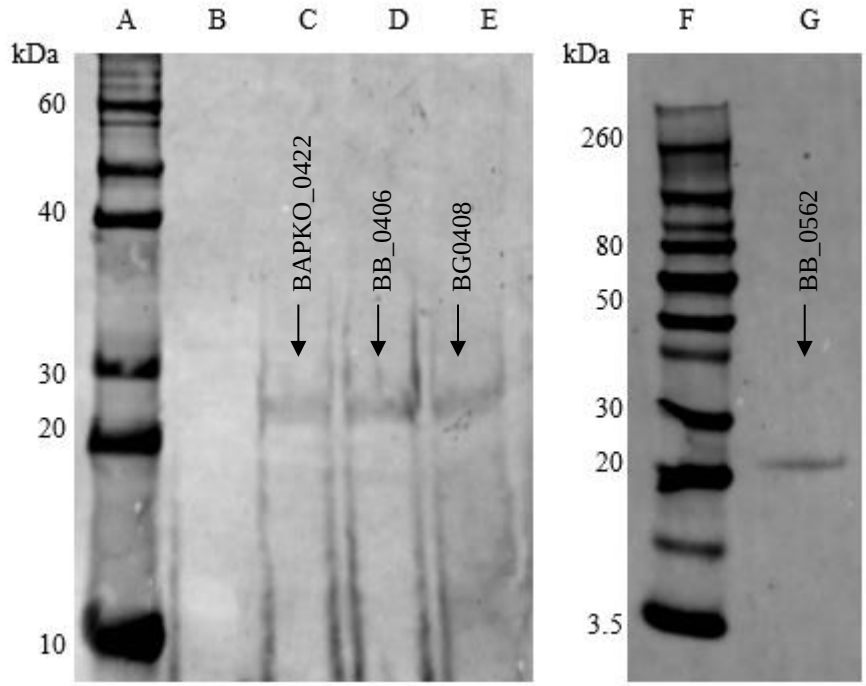


Figure 3.8. Western blot confirmation of *B. burgdorferi s.l.* Omp-A like proteins. Western blot analysis of cell lysates following IPTG controlled expression of *B. burgdorferi s.l.*OmpA-like proteins. **A)** Novex protein standard. **B)** Untagged lysozyme. **C)** BAPKO_0422, ~23kDa. **D)** BB_0406, ~23kDa. **E)** BG0408, ~23kDa. **F)** Novex protein standard. **G)** BB_0562, ~20kDa.

3.4 Refolding and purification of *B. burgdorferi s.l.* OmpA-like proteins

3.4.1 Preparation of inclusion bodies

One of very few positives of working with insoluble proteins was the ease in which a protein sample can be purified to a fairly high grade without employing any chromatography. Following lysis the samples were centrifuged and the supernatant or soluble fraction was discarded. The insoluble pellets containing the cell debris and inclusion bodies were prepared using a multistep wash procedure described in section 2.2.4.2. By the end of the wash stages the material was solubilised in denaturant (8M urea, 0.3M NaCl, 50mM tris, pH8) and the protein was approximately 80% pure for all four protein samples (BAPKO_0422, BB_0562, BB_0406, BG0408), example data shown for BAPKO_0422 in figure 3.9. Some level of target protein was lost during the pellet washing steps as observed by SDS-PAGE analysis of the wash supernatants shown in figure 3.9 however this loss was acceptable in order to improve the sample purity prior to refolding.

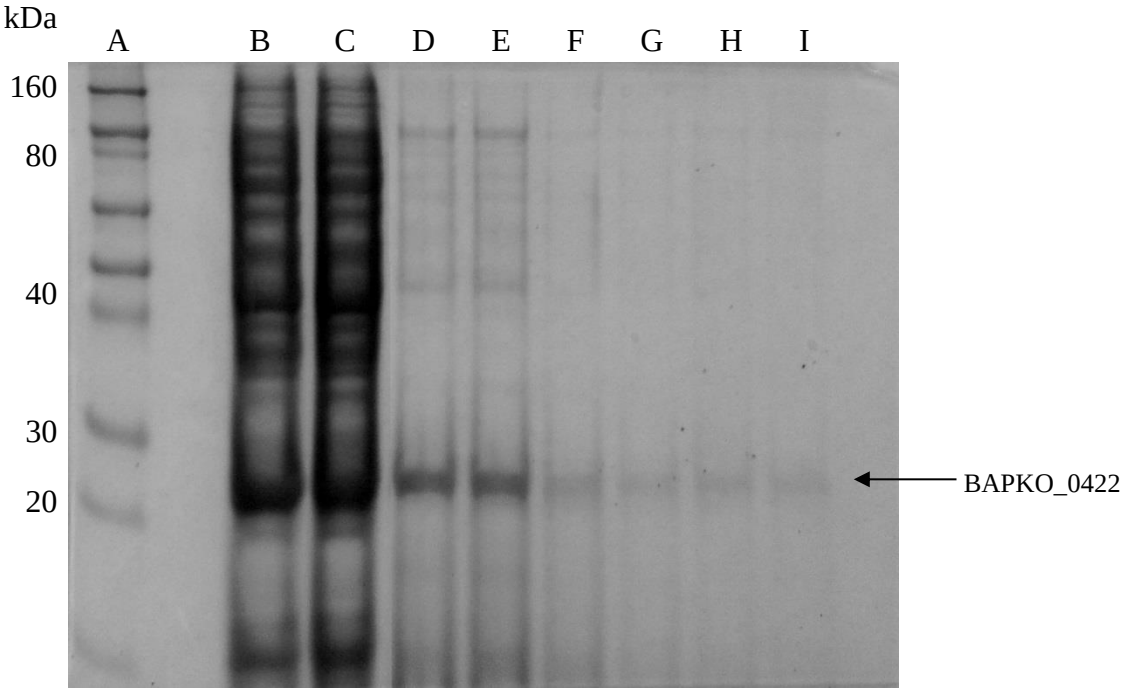


Figure 3.9. SDS-PAGE analysis of BAPKO_0422 following inclusion body preparation. Inclusion body pellets were subjected to extensive washing prior to solubilisation in denaturant. At each wash step a sample of the supernatant was retained for analysis where minimal BAPKO_0422 protein is lost. SDS-PAGE performed using pre-cast bis/tris 4-12% gels (Novex) and MES buffer (50mM MES, 50mM Tris base, 0.1% w/v SDS, 1mM EDTA, pH 7.3). **A)** Novex Protein Standard. **B)** BAPKO_0422 IPTG controlled insoluble fraction. **C)** BAPKO_0422 autoinduction insoluble fraction. **D)** BAPKO_0422 IPTG controlled following pellet washing. **E)** BAPKO_0422 autoinduction following pellet washing. **F-I)** Supernatants from the pellet washing.

3.4.2 IMAC purification of OmpA-like proteins

Early purification of BAPKO_0422 proved troublesome, several buffers and a range of pHs were trialled (section 2.2.5.1). However, the protein failed to bind to the Ni-NTA column, resulting in extremely poor yields and the majority of the protein remaining in the column flow through. Attempts were made to purify both under denaturing conditions in hope that the tag would be further exposed and also attempts were made to purify by Ni-NTA following refolding by rapid dilution. However, the binding yield remained low for BAPKO_0422 showed in figure 3.10. Further attempts were made to resolve the binding problem and involved the addition of glycerol (5% v/v) to the equilibration buffer and to the sample in hope of disrupting any hydrophobic interactions (Bane et al., 2007) that could be responsible, binding did improve slightly but the yield remained lower than expected. This poor binding to Ni-NTA columns was curious as the other his-tagged proteins including BB_0562, BB_0406 and BG_0408 were not exhibiting the same behaviour.

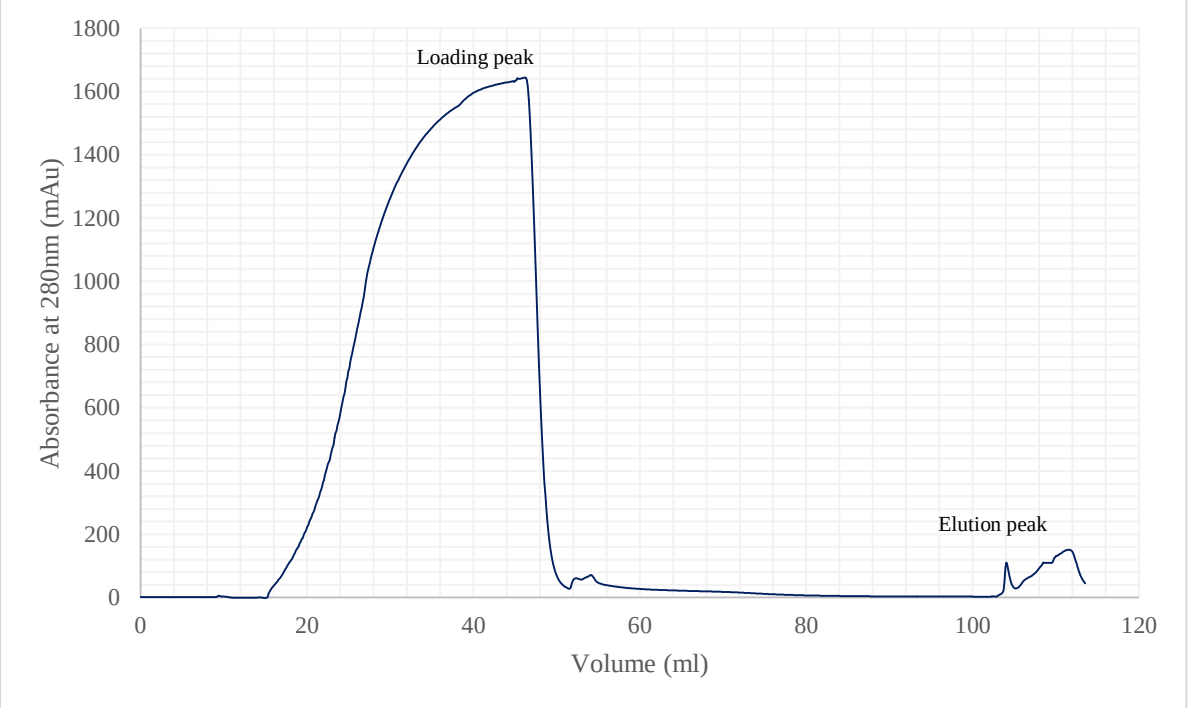


Figure 3.10. Low binding IMAC chromatogram for BAPKO_0422. Ni-NTA affinity chromatography of BAPKO_0422 showing poor yield. Approximately 35 ml of solubilised BAPKO_0422 (8M Urea, 0.3M NaCl, 50mM tris base, pH8) was pumped onto the pre-equilibrated His-trap column (GE Healthcare) and washed with 10 column volumes of 8M urea, 0.3M NaCl, 50mM tris, 20mM imidazole, pH8 prior to elution by a gradient increase of the imidazole concentration within the eluent (0.3M NaCl, 50mM tris, 0.3M imidazole, 0.1% v/v LDAO, pH8). The first peak at ~20-50 ml is the loading peak, the second peak at ~104 ml is the elution peak BAPKO_0422 confirmed by SDS-PAGE analysis) followed by third increase in absorbance due to the low grade imidazole used for elution.

Following a UV scan of the BAPKO_0422 protein sample in order to determine its concentration following pellet washing it was noted that the absorbance at 260 nm was particularly high possibly indicating DNA contamination. This was not originally suspected due to the expression cells undergoing extensive sonication during the lysis procedure. These suspicions that BAPKO_0422 samples were heavily contaminated with DNA were confirmed by ion exchange chromatography and the resulting fractions subjected to a UV scan data shown in figure 3.11. However after treatment of refolded BAPKO_0422 with DNase the binding of the protein to the nickel column increased over 10 fold, shown in figure 3.12. After discovery of the high DNA content of BAPKO_0422, all lysates for all of the OmpA-like proteins were subject to a DNase incubation following refolding by rapid dilution or IEC prior to IMAC on-column refolding and no further purification problems were encountered.

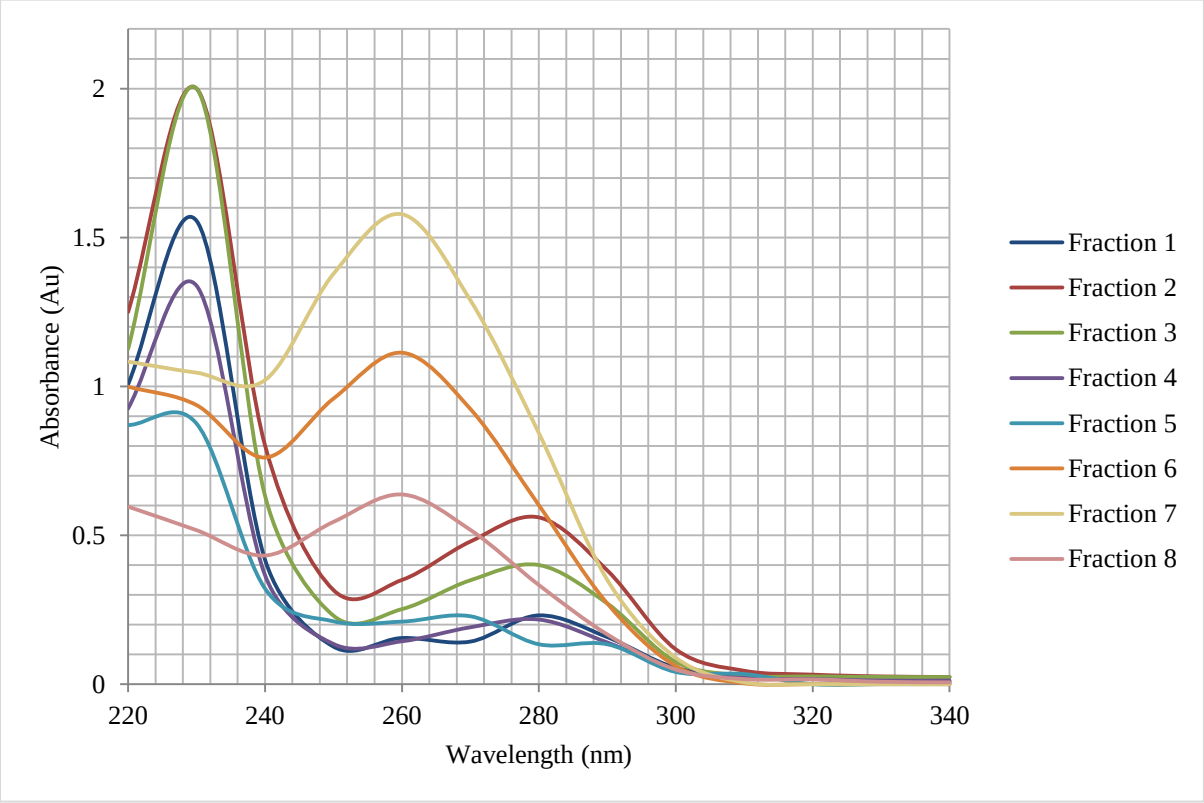


Figure 3.11. UV absorbance spectrum of BAPKO_0422 IEC fractions prior to refolding. BAPKO_0422 solubilised in urea was subjected to IEC in order to separate protein from DNA, fractions analysed by UV spectroscopy. Fractions 1-5 show a clear protein profile and were pooled and used for IMAC purification and on-column refolding. Fractions 6, 8 and 8 were indicative of high levels of DNA and discarded following SDS-PAGE analysis. Comparative spectral scans of DNA/Protein ratios can be seen in figure 2.4 of the methods section.

3.4.3 Refolding of BAPKO_0422 and BB_0562

Refolding of BAPKO_0422 and BB_0562 was less problematic than BB_0406 and BG0408, neither protein contained cysteine residues and the protein appeared to refold readily by removal of the denaturant and addition of the detergent LDAO. For both proteins this was achieved by both rapid dilution and by on-column refolding through a gradient removal of urea (section 2.2.5). On-column refolding was the preferred method as it avoided the processing of large volumes of protein solution that rapid dilution often generates. However, both proved to provide similar yields. As with all refolding strategies protein yield is low when compared to expression of soluble targets however both BAPKO_0422 and BB_0562 often delivered a final yield of 20-40% which was dependent upon the length and speed of the refolding gradient during on-column refolding. Yields were calculated based upon total protein and volume of the urea solubilised inclusion body fraction when compared to the volume and protein concentration of the refolded eluate following concentration and centrifugation to remove large aggregates.

3.4.4 Refolding of BB_0406 and BG_0408

Both BB_0406 and BG_0408 contained 2 cysteine residues which in both sequences these cysteines were separated by 8 residues. Due to presence of these cysteine residues the refolding yield was much lower than previously seen with BAPKO_0422 and BB_0562. Adaptations were made to protocols and buffer contents which included refolding under reducing conditions by rapid dilution and on-column refolding with the addition of further stabilisers such as glycerol (1-5% v/v) and L-arginine (0.1-0.5M). Protein yield remained fairly low following extensive optimisation and the original methods were put back into place. However, expression culture volumes were trebled in order to produce enough final product for downstream experiments. SDS-PAGE analysis of both BB_0406 and BG_0408 indicated that the recombinant protein could form oligomers including dimers, tetramers and octamers. These remained visible but at smaller proportions following reducing SDS-PAGE as shown in figure 3.13. These oligomers were confirmed by immunoblotting for the polyhistidine tag following separation by non-reducing SDS-PAGE.

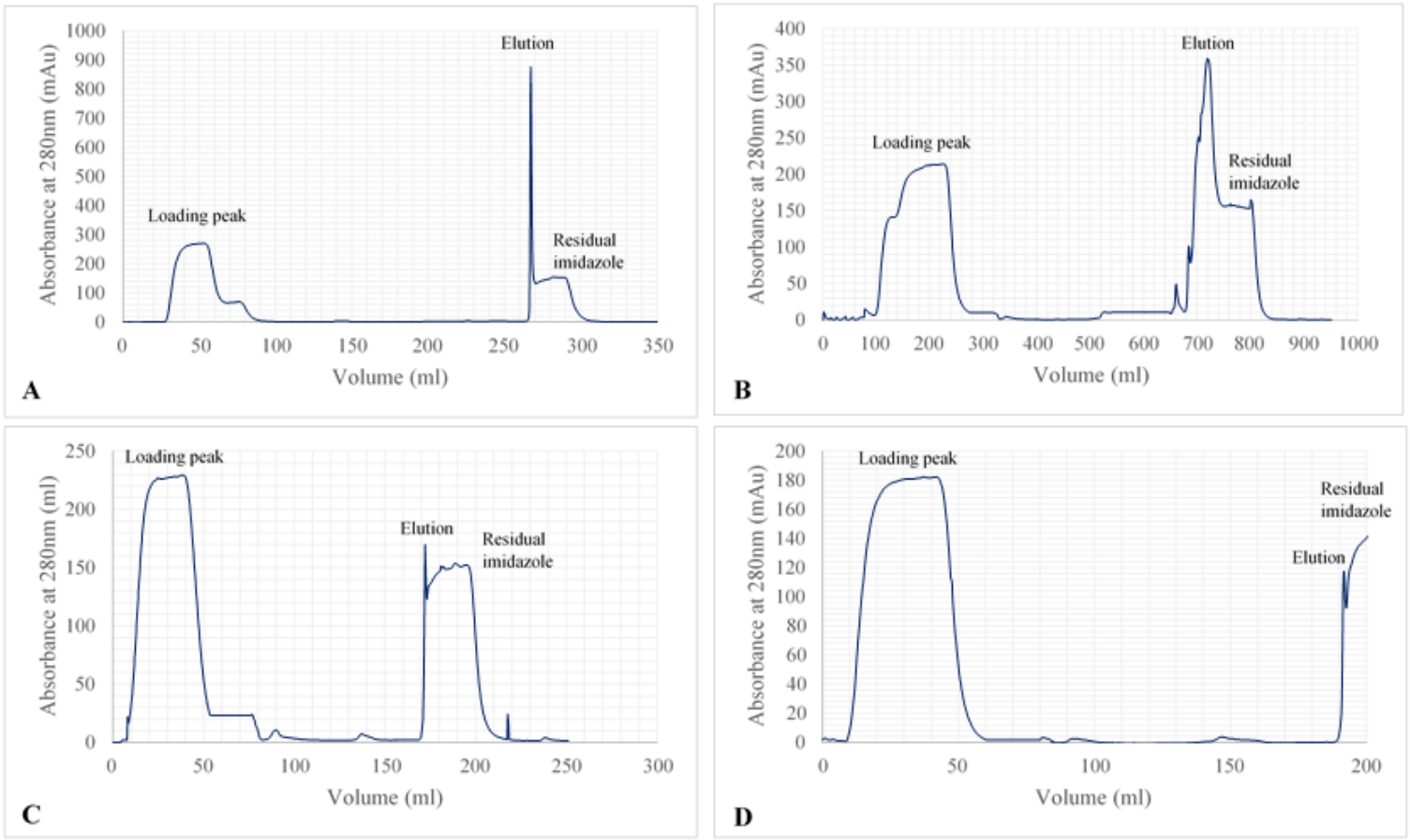


Figure 3.12. IMAC chromatograms for *B. burgdorferi s.l.*OmpA-like proteins. All IMAC purifications were undertaken using an Akta Prime chromatography system. Proteins were introduced to the pre-equilibrated column in 8M urea, 0.3M NaCl, 50mM tris, pH8 or for the case of BB_0562 without the denaturant urea. Following protein binding the column was washed with 10 column volumes of wash buffer (8M urea, 0.3M NaCl, 50mM tris, 50mM imidazole pH8) in order to remove non-specifically bound proteins. Where protein folding was achieved on the column the denaturant urea was removed by gradient reduction. Proteins were finally eluted from the column by a gradient increase of imidazole within the eluent (0.3M NaCl, 50mM tris, 0.3M imidazole, 0.1% v/v LDAO, pH8). A) BAPKO_0422 on-column refold and purification. B) BB_0562 extended purification following refolding by rapid dilution. C) BB_0406 on-column refold and purification. D) BG0406 on-column refold and purification.

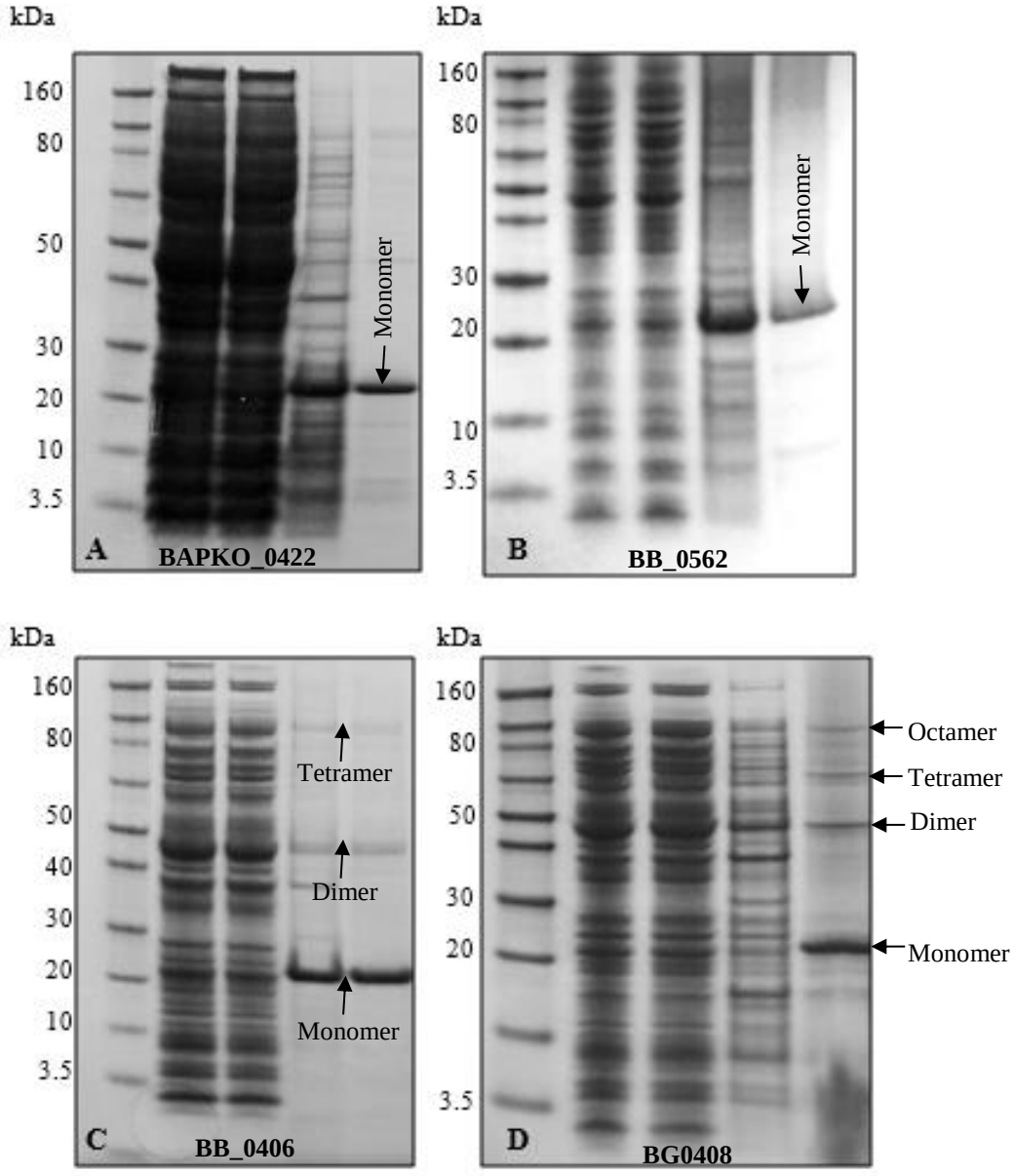


Figure 3.13. SDS-PAGE analysis of Omp-A like proteins following immobilised metal affinity chromatography. Samples were retained during the extraction and purification procedure and analysed by reducing SDS-PAGE. SDS-PAGE performed using pre-cast bis/tris 4-12% gels (Novex) and MES buffer (50mM MES, 50mM Tris base, 0.1% w/v SDS, 1mM EDTA, pH 7.3). **A)** Analysis of BAPKO_0422 samples. Lane 1 – Novex protein marker, 2 – Cell lysate, 3 – Soluble fraction, 4 – Insoluble fraction, 5 – Pooled eluate following Ni-NTA chromatography. **B)** Analysis of BB_0562 samples. Lane 1 – Novex protein standard, 2-3 Soluble fraction, 4 – Insoluble fraction, 5 – Pooled eluate following Ni-NTA chromatography. **C)** Analysis of BB_0406 samples. Lane 1 – Novex protein standard, 2-3 – Soluble fractions, 4 – Insoluble fraction, 5 – Pooled eluate following Ni-NTA chromatography. D) Analysis of BG0408 samples. Lane 1 – Novex protein marker, 2-3 Soluble fractions, 4 – Pellet wash fraction, 5 – Pooled eluate following Ni-NTA chromatography.

3.4.5 Gel filtration of the OmpA-like proteins

All OMPA-like proteins were subjected to gel filtration following the construction of a calibration curve shown in section 2.2.5.8. The column was equilibrated with an appropriate buffer matching the protein's buffering solution. Filtration was undertaken at a flow rate between 0.2 and 0.5 ml/min dependent upon the over column pressure detected for each sample. Protein samples were concentrated to 5 mg/ml and injected onto the top of the column. The UV absorption at 280 nm was monitored and fractions collected during the elution peak for SDS-PAGE analysis. Elution volumes were used to calculate the estimated molecular weight using the calibration curve previously described. Gel filtration chromatograms for OMPA-like proteins are shown in figure 3.14 with final elution volumes and molecular weight estimations shown within the figure caption.

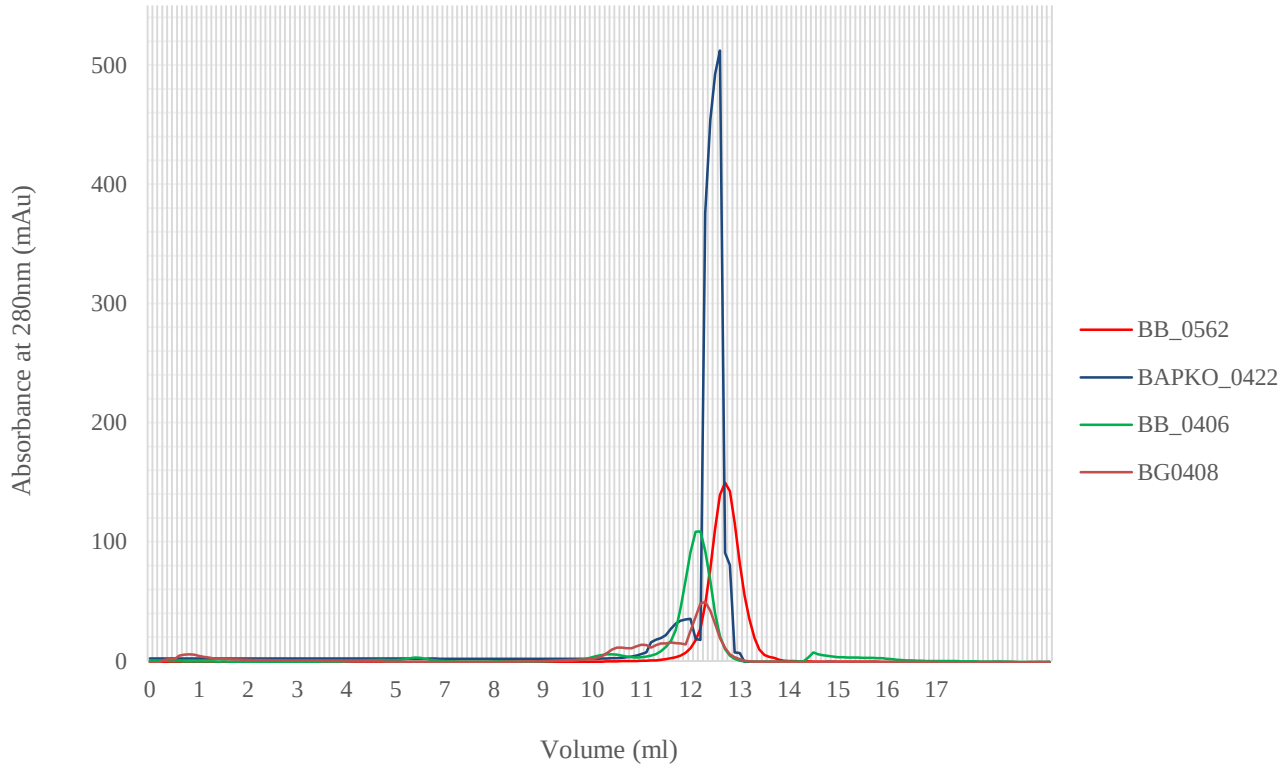


Figure 3.14. Size exclusion chromatography of *B. burgdorferi s.l.*OmpA-like proteins. Samples of all OmpA-like proteins were subjected to gel filtration in order to assess the refolded state, determine an approximate molecular weight and to remove any large aggregates prior to structural experiments. All gel filtration was performed under reducing conditions. The injection point for all samples is set at 0 ml whereby the protein sample was injected into the top of the superdex 75 10/300 GL column (GE Healthcare). All peak fractions were collected and analysed by SDS-PAGE. RED – BB_0562 at 2 mg/ml (~21kDa). BLUE – BAPKO_0422 at 5 mg/ml (~23kDa). GREEN – BB_0406 at 1.5 mg/ml (~22.5kDa). ORANGE – BG0408 at 2 mg/ml (22.5kDa).

For all four OMPA-like proteins samples were often found as polydisperse with some levels of aggregation and a proportion of monomer and oligomer. Monomeric fractions for each protein were pooled and concentrated. For the case of BB_0406 and BG_0408 a proportion of the protein was found to be dimeric, tetrameric and octameric although SEC was performed under reducing conditions (0.3M NaCl, 50mM Tris, 0.1% v/v LDAO, 1mM DTT, pH8). These oligomers did not appear as sharp peaks on the chromatogram shown in figure 3.14 but were found following SDS-PAGE of the eluate fractions before the peak representing the monomeric species. These fractions were removed with only monomeric species carried forward for structural experiments.

3.4.6 UV spectral analysis of *B. burgdorferi* OmpA-like proteins

Following early issues with regards to the presence of DNA within BAPKO_0422 samples all proteins samples were assessed by UV spectroscopy (section 2.2.5.7) following refolding and purification steps as shown in figure 3.15. All four OmpA-like protein samples contained minimal DNA contamination following the end stage of gel filtration.

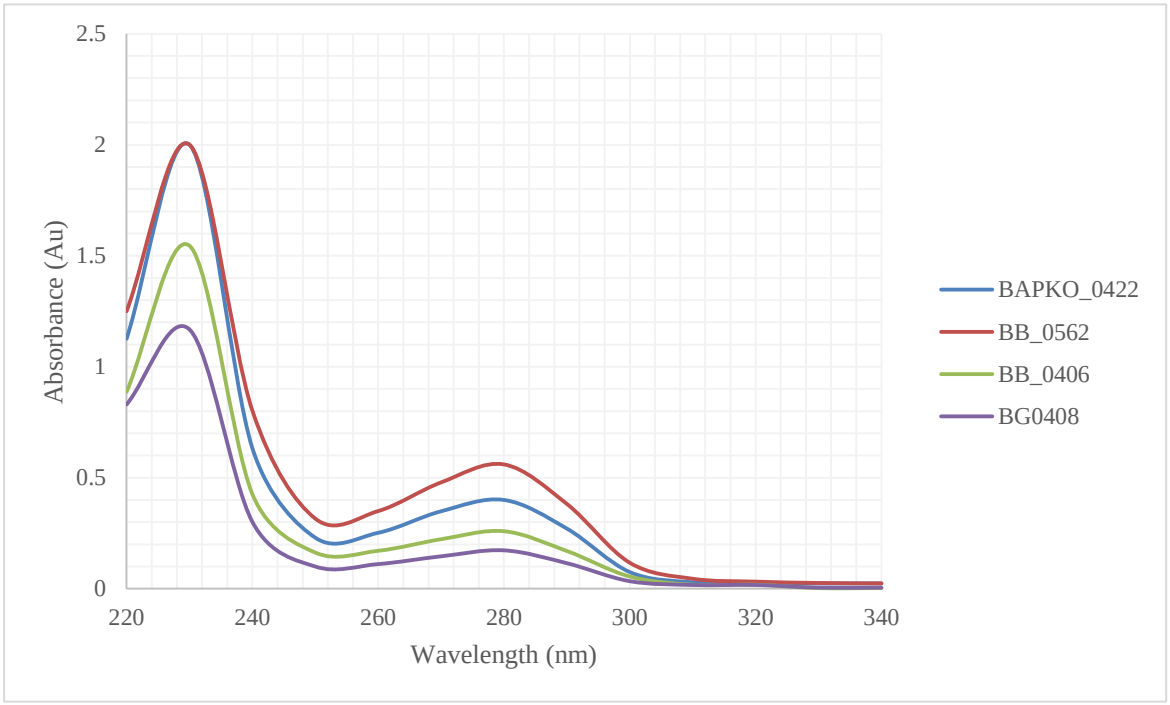


Figure 3.15. UV spectral analysis of *B. burgdorferi* OmpA-like proteins. Final UV spectrum for refolded and purified *B. burgdorferi* OmpA-like proteins. Scans were carried out on a UV Spectrostar Nano (BMG Labtech) with readings taken at 5 nm intervals from 220 to 340nm.

3.5 Gel shift assays for *B. burgdorferi* OmpA-like proteins

The formation of *E. coli*'s OmpA secondary structure and a number of other bacterial OMPs can be easily monitored by gel shift assays (section 2.2.6.1). These experiments exploit the characteristic that such proteins display a shift in molecular mass during SDS-PAGE. *E. coli* OmpA is seen to run at approximately 30kDa when the protein is fully folded but this value shifts upto 35kDa when the protein is either partially or fully unfolded. Although this phenomenon is seen for many β-barrel membrane proteins it was not observed for any of the four *B. burgdorferi* OMPs studied within this research (An example gel shift assay for BB_0562 is shown in figure 3.16). Although this differs from what is considered 'the norm' for this protein family it is not unheard of that some bacterial OMPs do not exhibit a gel shift and include *B. burgdorferi* P66 (Kenedy et al., 2014). As previously discussed the *Borrelia* porin P66 also does not display any difference in molecular weight when folded as opposed to its unfolded state. This lack of size shift for these OMPs could be a general characteristic of *B. burgdorferi* OMPs (Kenedy et al., 2014, Dyer et al., Appendix 4).

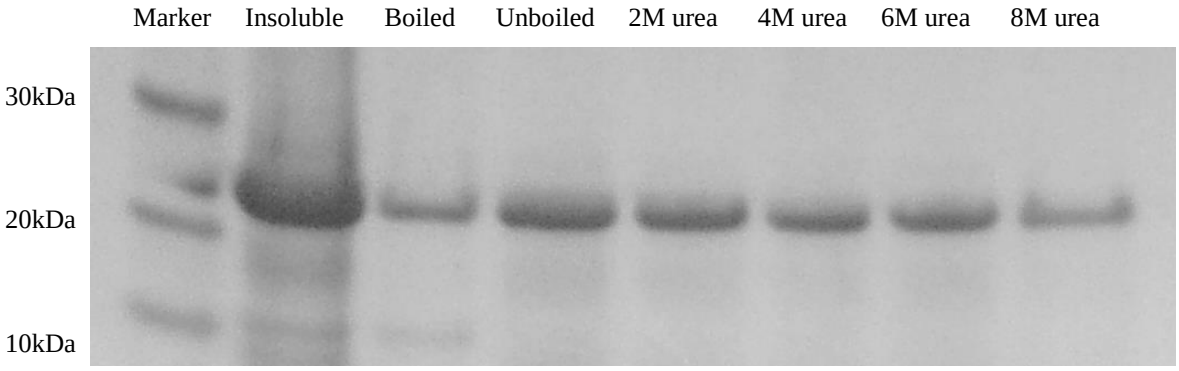


Figure 3.16. Gel shift assay for BB_0562. Samples of purified and refolded BB_0562 were denatured using varying concentrations of urea and analysed by SDS-PAGE. No shift in molecular weight was seen when comparing folded and unfolded samples. This behaviour was observed for all four *B. burgdorferi* OmpA-like proteins.

3.6 Phase partitioning of *B. burgdorferi* OmpA-like proteins

Phase partitioning experiments (section 2.2.6.2) were conducted in line with published methods (Bordier, 1981) for all four *Borrelia* OmpA-like proteins with data shown in figure 3.16. The low cloud point of Triton X-114 was exploited in order to separate amphiphilic membrane proteins from soluble proteins at a reasonably physiological temperature (~23ºC). Haemoglobin (Sigma) was selected as an aqueous negative control with *Arabidopsis thaliana* voltage-dependent anion channel 1 (AtVDAC1) selected as a positive control. Phase partitioning experiments demonstrated that all

four of the proposed *Borrelia* β-barrels appeared within the detergent rich phase indicating that they are all amphiphilic and highly likely to be integral membrane components.

BB_0406 and BG_0408 appeared in the detergent phase as monomeric, dimeric, tetrameric and a small proportion of octameric protein was present although unclear in figure 3.17 (blots E and F).

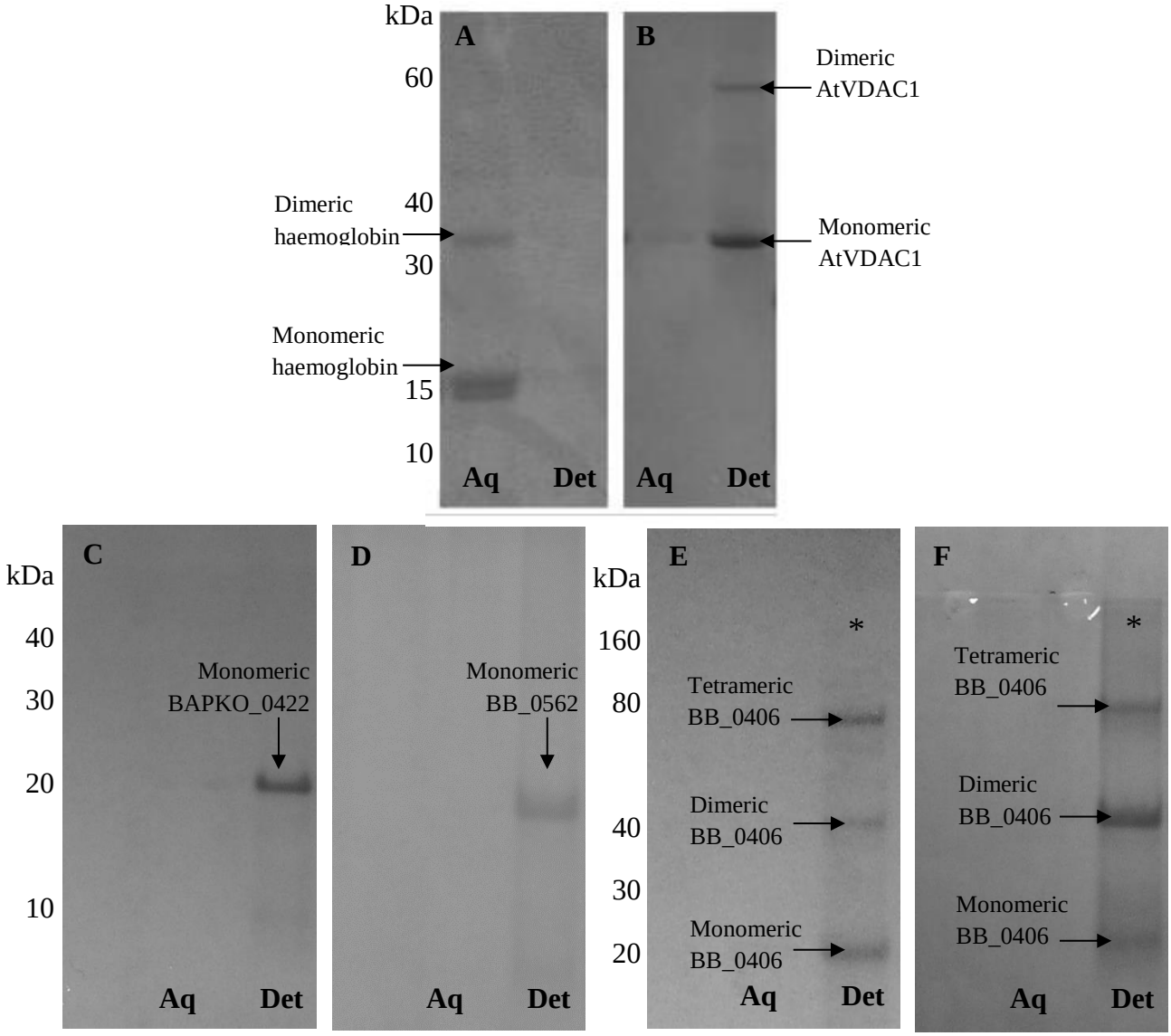


Figure 3.17. Phase partitioning of *Borrelia* outer membrane proteins. SDS-PAGE analysis following the Triton X-114 phase partitioning of possible *Borrelia* membrane proteins and controls. Lane 1 – aqueous phase, lane 2 – detergent phase. A) Negative control – Haemoglobin. B) Positive control – AtVDAC1, C) BAPKO_0422, D) BB_0562, E) BB_0406, F) BG0408.

3.7 Affinity ligand binding immunoblots (ALBI): Assaying for human factor H (hfH) binding

Initial hfH binding assays (section 2.2.6.3) were carried out for only for BAPKO_0422 and required much optimisation. The limiting factor of the experiment was the cost of purchasing natural hfH, initially it was hoped that the assay could be carried out with a hFh incubation concentration close to that found in human plasma, somewhere between 200-500 μg/ml, however this would have cost in the region of £1000 per incubation. In order to reduce the cost and to prevent a considerable reduction to the incubation concentration the assay was carried out using a minimal amount of PVDF membrane and the volume of the hFh for incubation decreased to 2.5 ml, giving a final incubation concentration of 77 μg/ml, approximately a third of what is found in human plasma. Initial practice assays carried out at this concentration and below produced a very weak positive signal and which was in some cases masked by poor non-specific binding due to issues with blocking. These problems were originally thought to be due to the low concentrations of hfH within the incubation. However, following optimisation it was noted that the chosen vessel to perform the assay had a slight curvature to the bottom resulting in an unequal wash of hfH over the PVDF membrane. By increasing the incubation volume from 2.5 ml to 3 ml these problems were resolved. Full assay optimisation can be seen in figure 3.18.

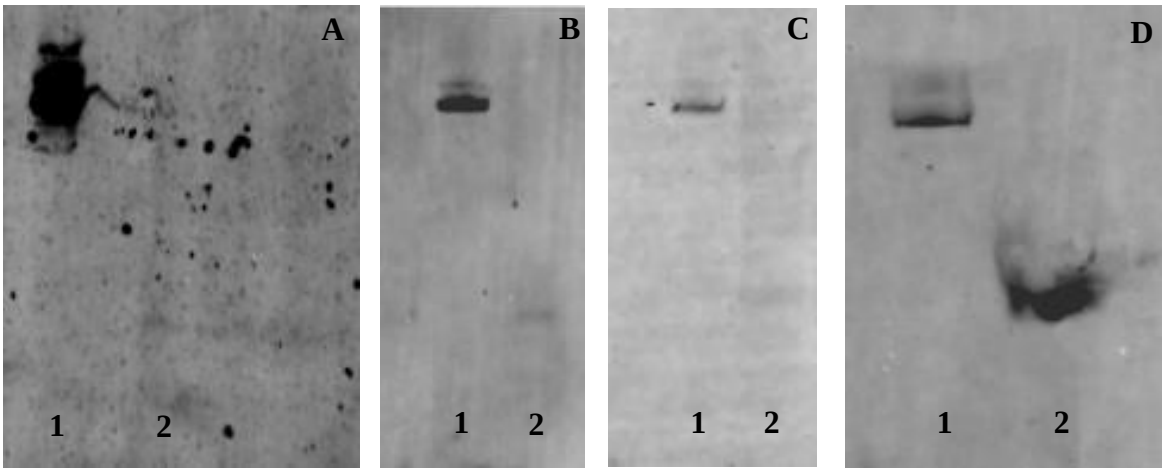


Figure 3.18. Optimisation of the hfH-BAPKO_0422 binding assay. Affinity ligand binding immunoblots of initial hfH-BAPKO_0422 assays. A) Initial experimental assay using 20μg of hfH as a positive control in lane 1 and 50μg of BAPKO_0422 in lane 2. B) Repeated assay with a weak positive signal. Lane 1 containing 0.3μg of hfH and 50μg of BAPKO_0422 in lane 2. C) Repeated assay with an increased amount of hfH used within the incubation, Lane 1 – 0.2μg of hfH, Lane 2 - 50μg of BAPKO_0422. D) Improved assay by increasing the incubation volume from 2.5 ml to 3 ml. Lane 1 – 0.2μg of hfH as a positive control, Lane 2 - 50μg of BAPKO_0422.

The second concern with the first hfH assays was the amount of hfH used as a positive control on the blot. Originally the amount of hfH used as the positive control was calculated based upon the average amount of protein required for effective transfer from an SDS-PAGE to a membrane, which is normally in the region of 20-30µg (Abcam, UK). However, this was far too high so further optimisation was required. An optimisation assay was carried out whereby various concentrations of hfH were subjected to native PAGE and subsequently transferred to PVDF which was blocked and probed with a primary anti-factor H (mouse anti-human - abcam) antibody followed by a secondary detection antibody (Alexa Fluor 680 goat anti-mouse), with data shown in figure 3.19.

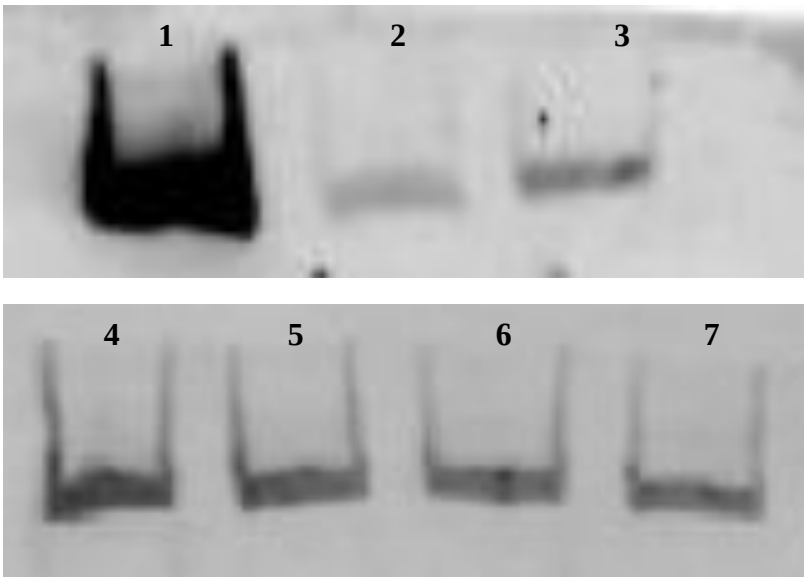


Figure 3.19. Optimisation of human fH as an ALBI positive control. Natural hfH run as a native PAGE and transferred to PVDF by wet transfer. Membranes were blocked with 5% w/v non-fat milk powder for 1hr, washed and probed with a primary anti-his followed by the secondary antibody**. Lanes** – (1) 0.5µg of hfH. (2) 0.1µg of hfH. (3) 0.25µg of hfH (4) 0.35µg hfH. (5) 0.3µg of hfH. (6) 0.25 µg of hfH. (7) 0.2µg of hfH.

Preceding the determination of an appropriate amount of hfH to use as a positive control several independent assays were undertaken for BAPKO_0422, all of which gave a positive result for hfH binding. Importantly a negative control blot was also carried out alongside every experimental assay. In this case the negative control followed the same process as of the experimental but the hfH incubation solution was substituted for TBS ensuring that neither antibody could give a false positive result. Extra proteins both natural and recombinant were also included on early blots as additional negative controls, these included BSA and the *Borrelia* protein SodA which was produced using the same expression system and carries the same N-terminal His-tag and HRV-3C recognition sequence.

Following extensive optimisation the final hfH-BAPKO_0422 binding assay was conducted with appropriate controls with full data shown in figure 3.20A. The ALBI assays resulted in a positive result for a hfH-BAPKO_0422 interaction. Following the positive assessment the assay was carried out for the remaining OMPs. BB_0562, BB_0406 and BG0408 with full data shown in figure 3.20A-D. All three of these protein appeared negative for hfH binding with the positive control strip of BAPKO_0422 again giving a positive result (Figure 3.21).

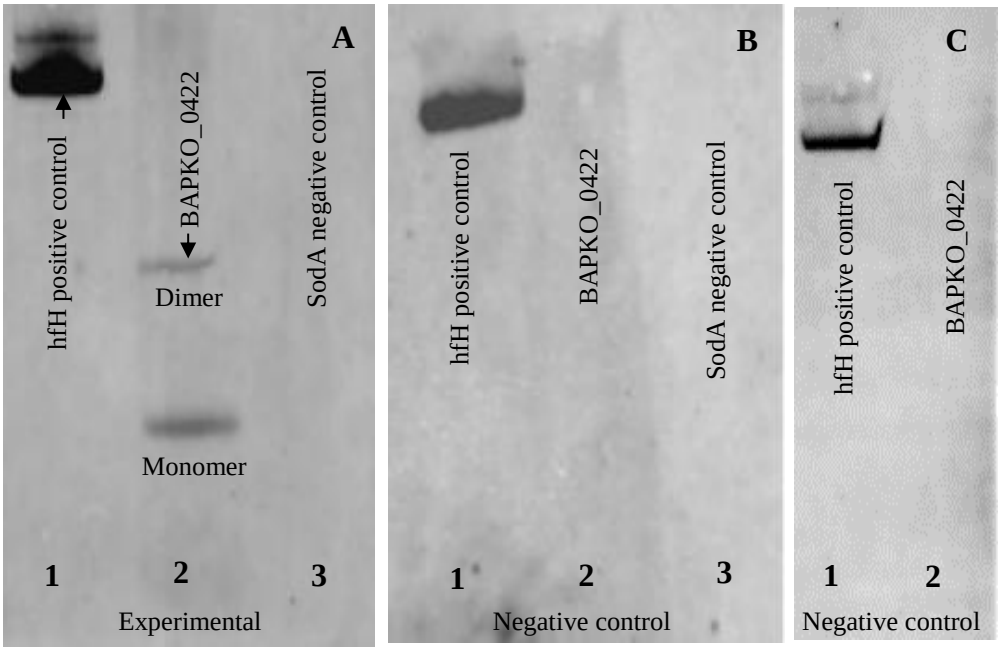


Figure 3.20. Assessing the interaction between hfH and BAPKO_0422. Affinity ligand binding immunoblots of hfH and BAPKO_0422 with control blots. **Blot A**. Experimental ALBI assay with a hfH incubation step. Lane 1 – 0.3µg of hfH. Lane 2 - 50µg of concentrated BAPKO_0422 showing a both monomeric and dimeric species. Lane 3 - 50µg of recombinant *B. burgdorferi* SodA. **Blot B**. Negative control ALBI assay carried out alongside blot A whereby the hfH incubation was substituted for a TBS incubation. Lane 1 - 0.3µg of hfH. Lane 2 - 50µg of BAPKO_0422. Lane 3 - 50µg of recombinant *B. burgdorferi* SodA. **Blot C**. Repeated negative control ALBI assay where the hfH incubation step was substituted with TBS. Lane 1 – 0.2µg of hfH. Lane 2 - 50µg of diluted BAPKO_0422.

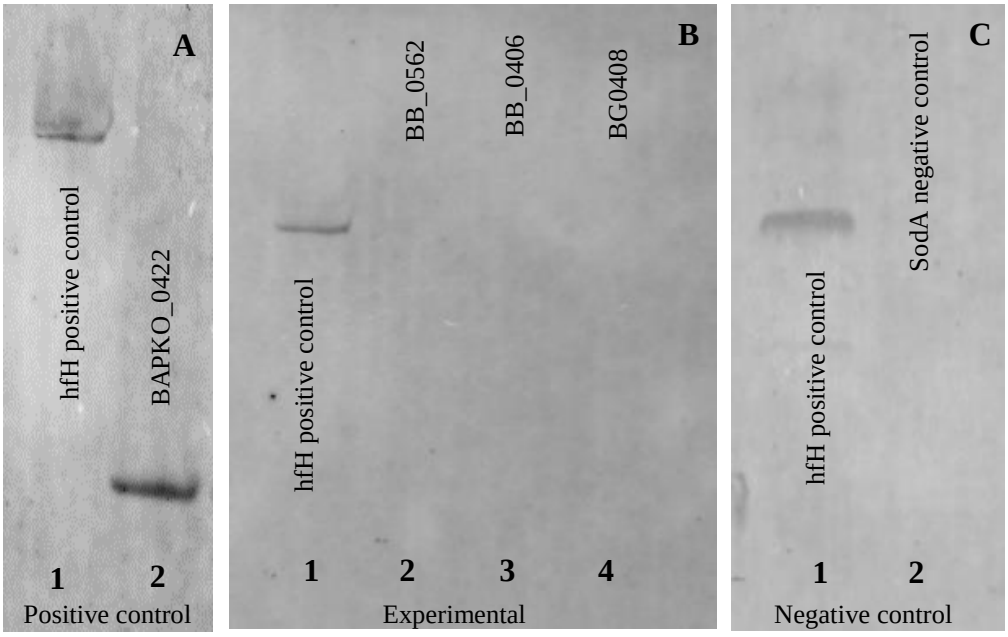


Figure 3.21. Final hfH ALBI assays for *Borrelia* OMPs. Affinity ligand binding blots were used to assess any possible interaction between hfH and the four possible OMP proteins. For BB_0562, BB_0406 and BG_0408 no interaction was observed. **Blot A**. Experimental ALBI assay for BAPKO_0422 with a hfH incubation step. Lane 1 – 0.2µg of hfH. Lane 2 - 50µg of diluted BAPKO_0422 showing a single monomeric species. **Blot B**. Experimental blot, Lane 1 – 0.2µg of hfH. Lane 2 - 50µg of BB_0562. Lane 3 - 50µg of BB_0406. Lane 4 - 50µg of BG0408. **Blot C**. Negative control blot. Lane 1 – 0.2µg of hfH. Lane 2 – 20µg of *B. burgdorferi* SodA.

3.8 BAPKO_0422 - Heparin binding assays

As previously discussed in section 3.4.2 recombinant BAPKO_0422 was seen to display a remarkable affinity for DNA which resulted in many problems during downstream chromatography experiments. BAPKO_0422 could only be purified by Ni-IMAC once the contaminating DNA was removed from the sample. These issues were not seen for the other three *Borrelia* OMPA-like proteins and raised the question whether BAPKO_0422 may interact with other negatively charged molecules such as the highly sulphated glycosaminoglycan, heparin. Although many proteins are known to interact with heparin or heparan sulphate, BAPKO_0422's possible outer surface location could suggest an important interaction.

The ability of BAPKO_0422 to bind to heparin was assessed as previously discussed in section 2.2.6.4. Recombinant BAPKO_0422 was fed onto a GE Healthcare Heparin column using an Akta Prime chromatography system and the resulting flowthrough and eluate was analysed using SDS-PAGE analysis with the full data shown in figure 3.22. Following the introduction of 5 ml of BAPKO_0422 at a concentration of 0.5 mg/ml to the heparin column no substantial level of protein was seen in either the low NaCl wash (50mM NaCl) or within the flowthrough nor detected by SDS-PAGE.

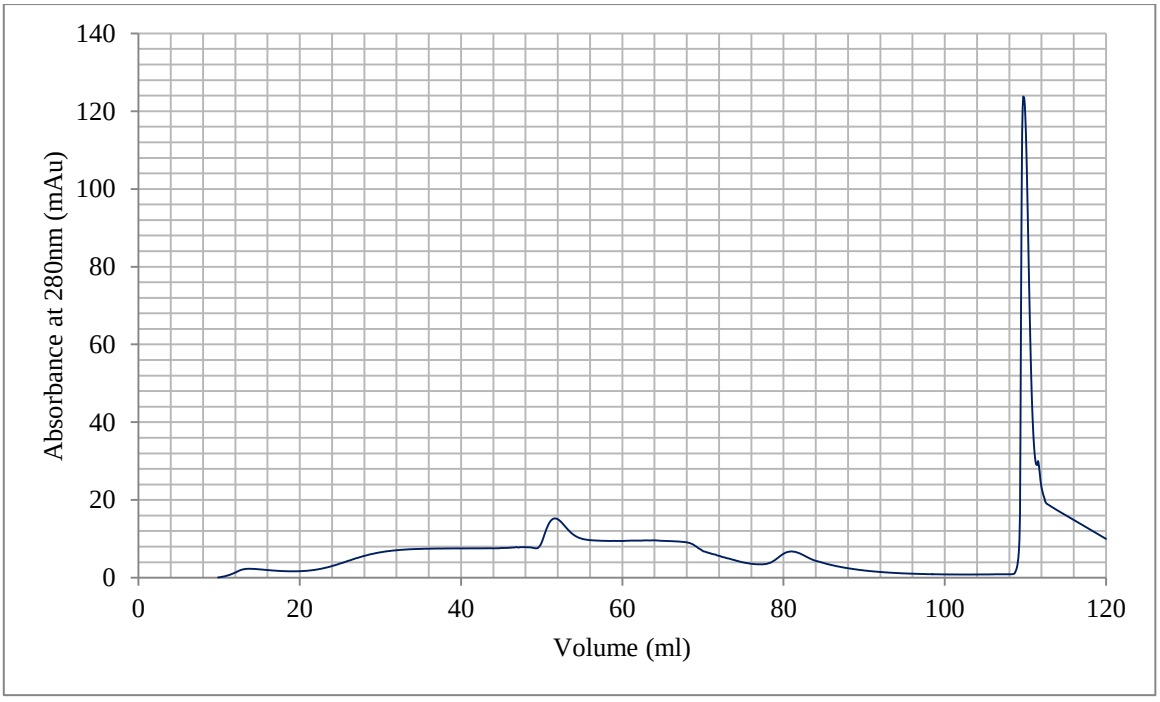


Figure 3.22. BAPKO_0422 heparin binding assay. Recombinant BAPKO_0422 was applied to a heparin column and quantitatively retained and eluted using 1M NaCl. No significant peak was observed during the loading or during the low NaCl wash (50mM NaCl) and the majority released from the column during elution.

The ability of BAPKO_0422 to bind heparin *in vitro* requires much further investigation in order to determine if this interaction is physiologically important *in vivo* during *B. burgdorferi* infection. The preliminary results require further confirmatory experiments including, Heparin competitive binding assays (Dawes & Pepper, 1987) and *in vivo* studies including knock out models (Snäll et al., 2016).

3.9. Circular dichroism experiments for *Borrelia* OMPs

3.9.1. Introduction to circular dichroism

Circular dichroism or CD is a common spectroscopic technique used for rapid assessment of a samples secondary structure. In technical terms CD can be defined as the disproportionate absorption of left and right-handed circularly polarised light (L-CPL and R-CPL) which occurs when the sample molecule contains at least one chiral chromophore.

A beam of light can be described to have both time dependent electric and magnetic fields. Should this beam become linearly polarised by the use of filters or prisms the beam is observed to oscillate within a single plane. All types of polarised light are described as the sum of two independent linearly polarised states, for example vertical and horizontal polarised light. When these vertical and horizontal polarised states are both in phase and of equal amplitude the resulting wave can be described as linearly polarised at 45 degrees but should vertical and horizontal waves become out of phase the resulting wave is no longer linearly polarised (Greenfield, 2006). In the case of CD should the two polarised states become out of phase by a quarter of a wave the resulting wave can represented as a helix or alternatively known as circularly polarised light (CPL) which can be either right or left handed and are non-superimposable mirror images (Greenfield, 2006). This generated sinusoidal wave can be explained by two vectors of equal length which rotate in a circular pattern, one clockwise (E_R) and the other counter-clockwise (E_L) and are 90 degrees out of phase (Greenfield, 2006). Asymmetric molecules which contain at least one chiral chromophores may absorb left and right handed circularly polarised light disproportionately. This results in a rotation in the plane of light and the addition of the E_R and E_L vectors results in a wave which now traces an elliptical shape, known as elliptically polarized light.

The CD of a sample is normally described as the difference in absorption of E_R and E_L known as ΔE shown in equation 4 or alternatively measured in degrees ellipticity which can be defined as the tangent of the ratio of the minor to major elliptical axis or molar ellipticity whereby the CD is corrected for sample concentration. CD units are often further converted to mean residue ellipticity a normalised unit which is particularly useful when using methods for estimating the secondary structure of proteins and is simply the molar ellipticity divided by the number of residues or amino acids within the molecule (Greenfield, 2006).

For polypeptides and proteins CD is a useful tool for the rapid determination of the folded state and for estimating the total amounts of secondary structure. As the amide groups of the polypeptide

backbone are chromophores and their optical transitions can be further split to multiple transmissions due to exciton interactions (Sreerama and Woody, 2000) typical structural elements of proteins give rise to specific CD spectra. These indicative spectra are explained by the differential absorption of the amide chromophores or peptide bonds when present in particular chiral environments created by protein folding (Kelly et al., 2005). An example of this phenomenon is seen with α-helical proteins which have defined negative bands at both 222 and 208 nm and a positive band at 193 nm (Holwarth & Doty, 1966) in contrast proteins containing high levels of antiparallel β-sheets have clear negative bands at 218 nm and a positive band at 195 nm (Greenfield & Fasman, 1969).

$$\textbf{Circular dichroism} = \Delta A(\lambda) = A(\lambda)_{LCPL} - A(\lambda)_{RCPL}$$

Equation 4. Calculation of CD. *Where λ is the wavelength*

3.9.2 Circular dichroism of *Borrelia* OMPs

CD data for all four proposed OMPs from *Borrelia burgdorferi, afzelii* and *garinii* were collected as previously described in section 2.2.8.1 using a JASCO J-810 spectropolarimeter (University of York). The resulting spectra are shown in figure 3.24. Recombinant OMPs were refolded, purified and centrifuged prior to data collection. Data were collected in quintuplicate across several concentrations for each protein sample in an effort to obtain the best spectra across the largest range of wavelength (full data are included within the appendix section 2). CD spectropolarimeters maintain a constant current by raising the voltage as the level of light falls. When scanning at lower wavelengths the absorbance increase causes the H[T] voltage to rise. This rise in voltage changes the ratio of signal to noise and at above 500 V data often becomes noisy and unreliable. This H[T] value is also known as the photomultiplier tube (PMT) voltage or dynode voltage with the cut-off value differing for each machine set up. Data collected preceding 600V H[T] were discarded due to reliability. All spectra were subjected to background subtraction using an appropriate blank with the resulting file used for deconvolution using the Dichroweb server (Whitmore & Wallace, 2007, Whitmore & Wallace, 2004, Lobley et al., 2002).

Early assessment of BAPKO_0422 by circular dichroism produced a spectrum which appeared shifted towards the far UV and although the characteristic beta-sheet curve was present as shown in figure 3.23 the positioning of the curve did not agree with data for other predominantly beta sheet proteins. It was hypothesised that it was likely that the sample contained a mixed species of folded and partially folded protein and the data were discarded. A second sample of BAPKO_0422 was

subjected to CD which gave rise to a more accurate curve as shown in figure 3.24. This data did not appear to hold the same shift towards far UV and was used for deconvolution using the dichroweb server. It was later found that the shift towards the far UV was caused by technical problems with the spectropolarimeter whereby changes in wavelength were not correctly recorded by the software.

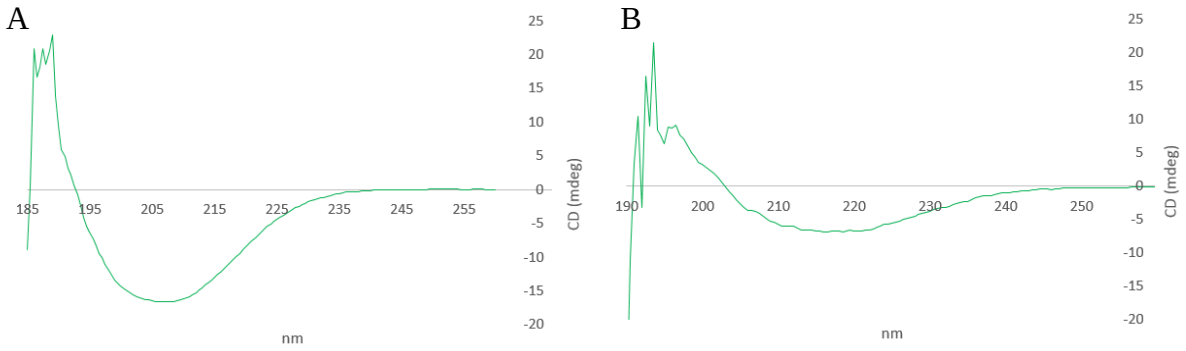


Figure 3.23. Circular dichroism spectra of BAPKO_0422 – **A**. CD spectra of BAPKO_0422 showing a shift of the curve to far UV due to issues with the JASCO spectropolarimeter. Spectra obtained for 0.1mg/ml BAPKO_0422 in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. **B**. CD spectra of BAPKO_0422 following the resolution of the problem. The H[T] value was reached at 195.5 nm and deconvolution calculated the secondary structure as 40% β-sheet, 4% helix, 20% turn and 35% unordered. Spectra obtained for 0.075mg/ml BAPKO_0422 in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8.

The H[T] value was reached at 195.5 nm and the data from 260 nm and down to this point was used for the estimation of the secondary structure. The sum of secondary structure was calculated to be 40% β-sheet, 4% helix, 20% turn and 35% unordered. Using the CDSSTR algorithm (Compton & Johnson, 1986, Manavalan & Johnson, 1987) and both SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011) data sets which were chosen as these data sets contain the largest number of reference proteins. The full contents of each data set can be found within appendix 2.

Following the initial analysis of BAPKO_0422 further samples were produced alongside BB_0562, BB_0406 and BG0408 for CD analysis. During these experiments the limiting value of H[T] was reached at longer wavelength than what was observed for the previous BAPKO_0422 experiments. Initially deconvolution was not going to be undertaken as data were needed down to 195 nm. However, the H[T] value was reached at around 200 nm leaving the data from 200 nm to 195 nm unreliable. Deconvolution was tested for BAPKO_0422 as the secondary structure had already been

estimated from the earlier data set. Both deconvolution data sets agreed well even with the inclusion of the 5 nm of unreliable data. Therefore deconvolution was undertaken for the remaining OMPs down to 195 nm which in some cases extended slightly past the acceptable H[T] value.

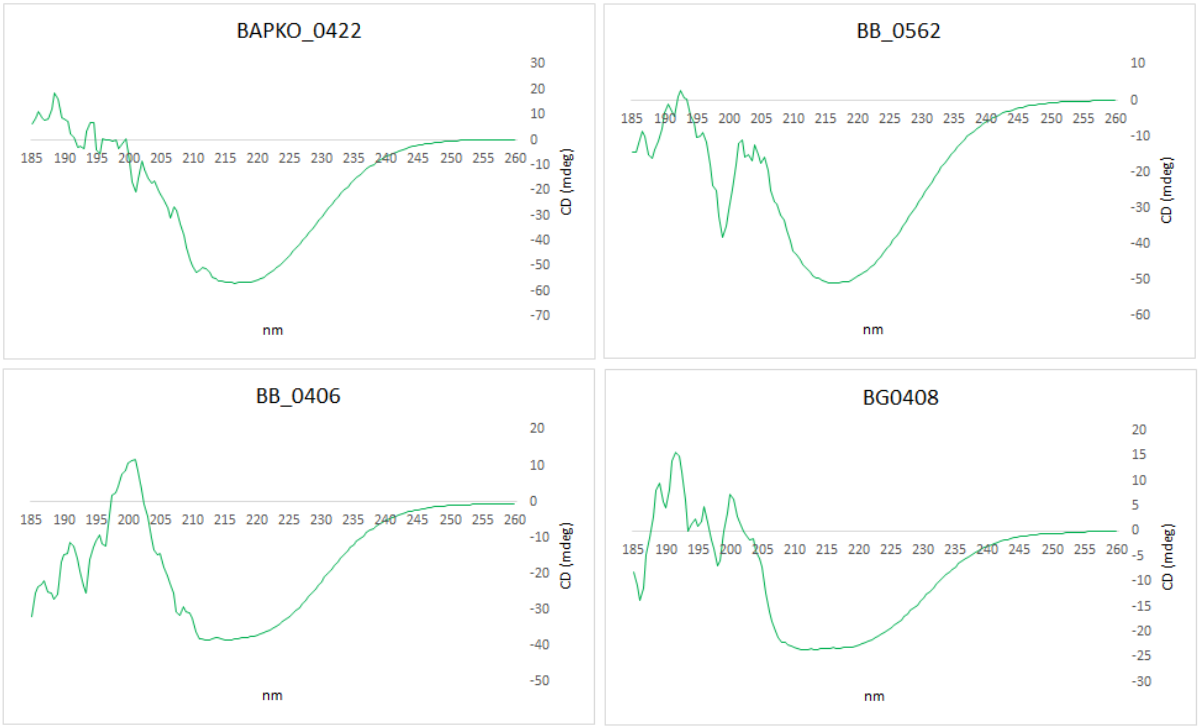


Figure 3.24. Circular dichroism spectra of *Borrelia* OMPs. The CD spectra for four proposed *Borrelia* OMPs. CD measurements were made on a JASCO J-810 with the measurements made from 260 nm to 185 nm at 0.5 nm increments. The spectra show all data including that once the H[T] reached 600V. Spectra for all proteins was obtained in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. BAPKO_0422 – Sample concentration of 0.25 mg/ml With H[T] reaching 600 V at 200 nm. BB_0562 – Sample concentration of 0.25 mg/ml with H[T] reaching 600V at 205 nm. BB_0406 – Sample concentration of 0.2 mg/ml with H[T] reaching 600V at 199 nm. BG0408 – Sample concentration of 0.125 mg/ml with H[T] reaching 600V at 200 nm.

Parameter	BAPKO_0422	BB_0562	BB_0406	BG0408
Experimental temperature	20ºC	20ºC	20ºC	20ºC
Path length	0.2cm	0.2cm	0.2cm	0.2cm
Protein concentration	0.25 mg/ml	0.25 mg/ml	0.2 mg/ml	0.2 mg/ml
File format	JASCO 1.30	JASCO 1.30	JASCO 1.30	JASCO 1.30
Wavelength range	260-185nm	260-185nm	260-185nm	260-185nm
Wavelength increments	0.5nm	0.5nm	0.5nm	0.5nm
Analysis algorithm	CDSSTR	CDSSTR	CDSSTR	CDSSTR
Reference set	SP175/SMP180	SP175/SMP180	SP175/SMP180	SP175/SMP180
Initial wavelength	260nm	260nm	260nm	260nm
Final wavelength	185nm	185nm	185nm	185nm
Input units	Mdeg	Mdeg	Mdeg	Mdeg
Output units	Mean residue ellipticity	Mean residue ellipticity	Mean residue ellipticity	Mean residue ellipticity
Mean residue weight (Da)*	111.9	110.2	112.8	112.8

**Table 3.3. Parameters selected for circular dichroism experiments for *Borrelia* OmpA-like proteins. ***The mean residue weight was calculated as the average molecular weight of all the amino acids within the protein sequence. The value is calculated by dividing the total molecular mass of the polypeptide by the number of residues within the sequence.

All four of the proposed *Borrelia* OMPs showed a spectrum typical of proteins containing large amounts of beta sheet with the distinctive minima close to 218 nm and maxima around 195 nm (Greenfield and Fasman, 1969) although in some cases this maxima was missed by noise due to the H[T] threshold being met.

Deconvolution for all four *Borrelia* OMPs was consistent when using two independent data sets and all four proteins appear to share a similar proportion of secondary structure. The deconvolution data is shown in table 3.14 and the experimental and reconstructed fit can be seen in figure 3.25. Deconvolution for each sample using both SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011) reference sets were averaged to give a final estimation of the secondary structure as summarised in table 3.4. Overall the four proteins appear to share a similar level of secondary structure and the CD experiments support that the four *Borrelia* proteins contain a high level of β-sheet (~42%) indicating that they could indeed be integral β-barrel membrane proteins.

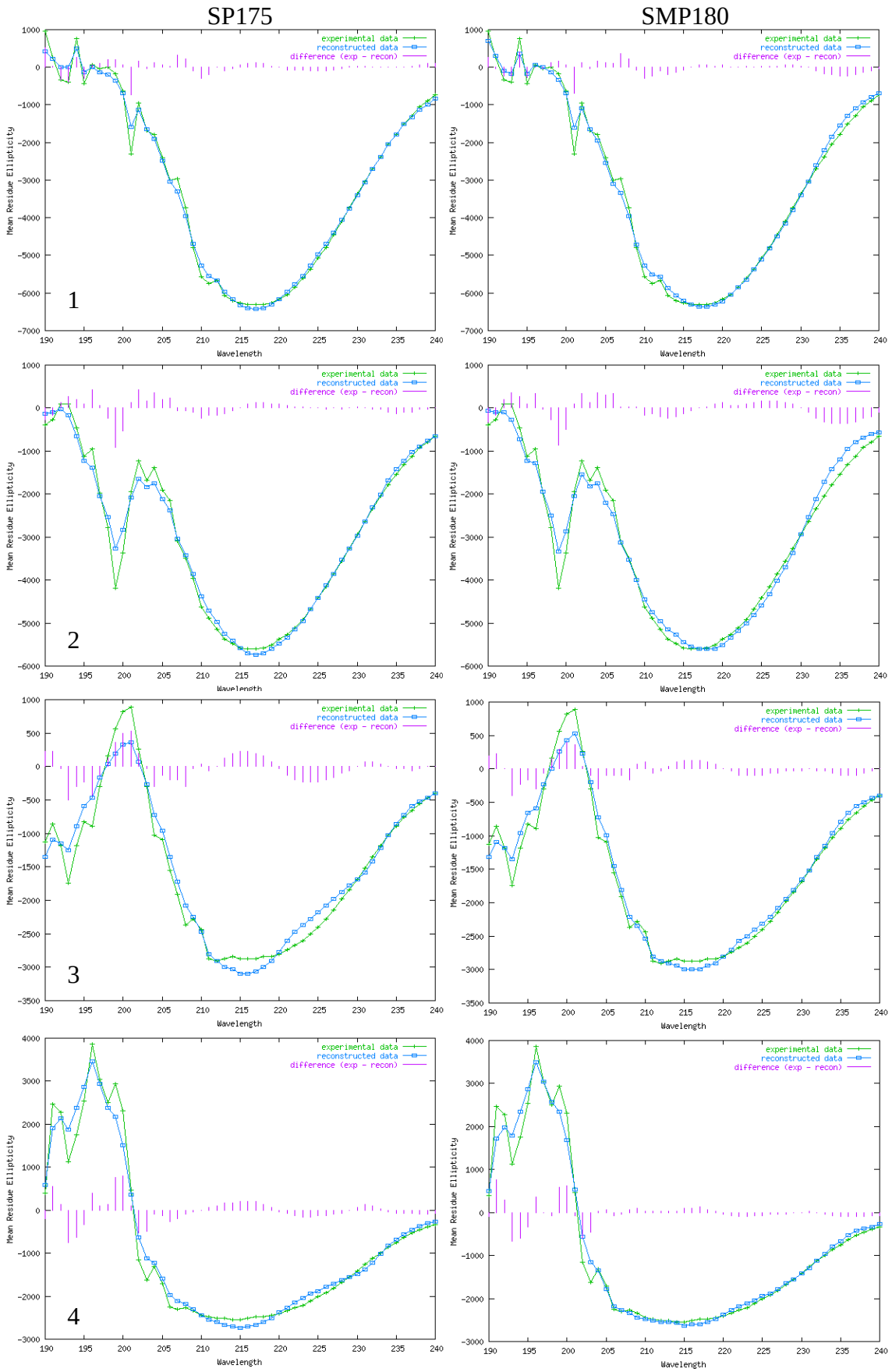


Figure 3.25. Deconvolution analysis of *Borrelia* OMPs. Plots were generated using the Dichroweb server using the CDSSTR algorithm (Compton & Johnson, 1986, Manavalan & Johnson, 1987) and SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011) as reference sets. (Continued…)

The left plots were produced using SP175 (Lees et al., 2006) and the right side plots were generated using SMP180 (Abdul-Gader et al., 2011) as the reference set. The green line represents the experimental data, the blue line is the reconstructed data using the CDSSTR (Compton & Johnson, 1986, Manavalan & Johnson, 1987) algorithm and the pink lines indicate the difference between the two. Areas where the pink line is at it's longest suggest a poor fit between the experimental and reconstructed data (Whitmore and Wallace, 2004). From top to bottom – **1.** BAPKO_0422, **2.** BB_0562, **3.** BB_0406, **4.** BG0408.

Sample	Reference set	Helix	Strand	Turns	Unordered	Total
BAPKO_0422 sample 1	SP175	0.04	0.42	0.12	0.4	0.98
BAPKO_0422 sample 1	SMP180	0.03	0.43	0.11	0.41	0.98
BAPKO_0422 sample 2	SP175	0.05	0.39	0.14	0.39	0.97
BAPKO_0422 sample 2	SMP180	0.03	0.47	0.11	0.38	0.99
BAPKO_0422 sample 3	SP175	0.05	0.40	0.13	0.41	0.99
BAPKO_0422 sample 3	SMP180	0.05	0.41	0.12	0.41	0.99
BB_0562 sample 1	SP175	0.06	0.41	0.12	0.40	0.99
BB_0562 sample 1	SMP180	0.03	0.44	0.11	0.40	0.98
BB_0562 sample 2	SP175	0.04	0.40	0.12	0.41	0.97
BB_0562 sample 2	SMP180	0.04	0.43	0.12	0.40	0.99
BB_0562 sample 3	SP175	0.04	0.40	0.12	0.42	0.98
BB_0562 sample 3	SMP180	0.04	0.42	0.12	0.40	0.98
BB_0406 sample 1	SP175	0.04	0.42	0.13	0.40	0.99
BB_0406 sample 1	SMP180	0.04	0.44	0.12	0.40	1
BB_0406 sample 2	SP175	0.03	0.43	0.12	0.40	0.98
BB_0406 sample 2	SMP180	0.03	0.44	0.12	0.39	0.98
BB_0406 sample 3	SP175	0.04	0.42	0.13	0.40	0.99
BB_0406 sample 3	SMP180	0.04	0.44	0.12	0.40	1
BG0408 sample 1	SP175	0.04	0.44	0.12	0.39	0.99
BG0408 sample 1	SMP180	0.03	0.44	0.14	0.38	0.99
BG0408 sample 2	SP175	0.06	0.41	0.13	0.40	1
BG0408 sample 2	SMP180	0.06	0.40	0.12	0.42	1
BG0408 sample 3	SP175	0.04	0.41	0.11	0.42	0.98
BG0408 sample 3	SMP180	0.04	0.40	0.11	0.43	0.98

Table 3.4. Deconvolution data for *Borrelia* OMPs. Deconvolution was carried out using the Dichroweb server based upon the CDSSTR algorithm (Compton & Johnson, 1986, Manavalan & Johnson, 1987) using both SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011) as data sets. Three different concentrations for each recombinant OMP were used.

Following analysis of the *Borrelia* CD data, spectra for structurally characterised 8-stranded β-barrels were generated using Dichrocalc. Theoretical CD data were generated using pdb files for *E. coli* OmpA (1BXW - Pautsch and Schulz, 1998), OmpW (2FIV - Hong et al., 2006) and OmpX (1QJ8 - Vogt & Schulz, 1999) alongside *Y. pestis* Ail (3QRA - Yamashita et al., 2011) and *N. meningitidis* NspA (1P4T - Vandeputte-Rutten et al., 2003) in order to compare the experimental *Borrelia* OMP data with data for structurally characterised membrane proteins as shown in figure 3.26. The predicted

CD spectra followed the same distinct shape as what was seen for the *Borrelia* proteins with a the minima being around 218 nm and the maximum reaching around 195 nm giving further support that the *Borrelia* proteins follow a similar CD profile to that of known 8-stranded β-barrel proteins.

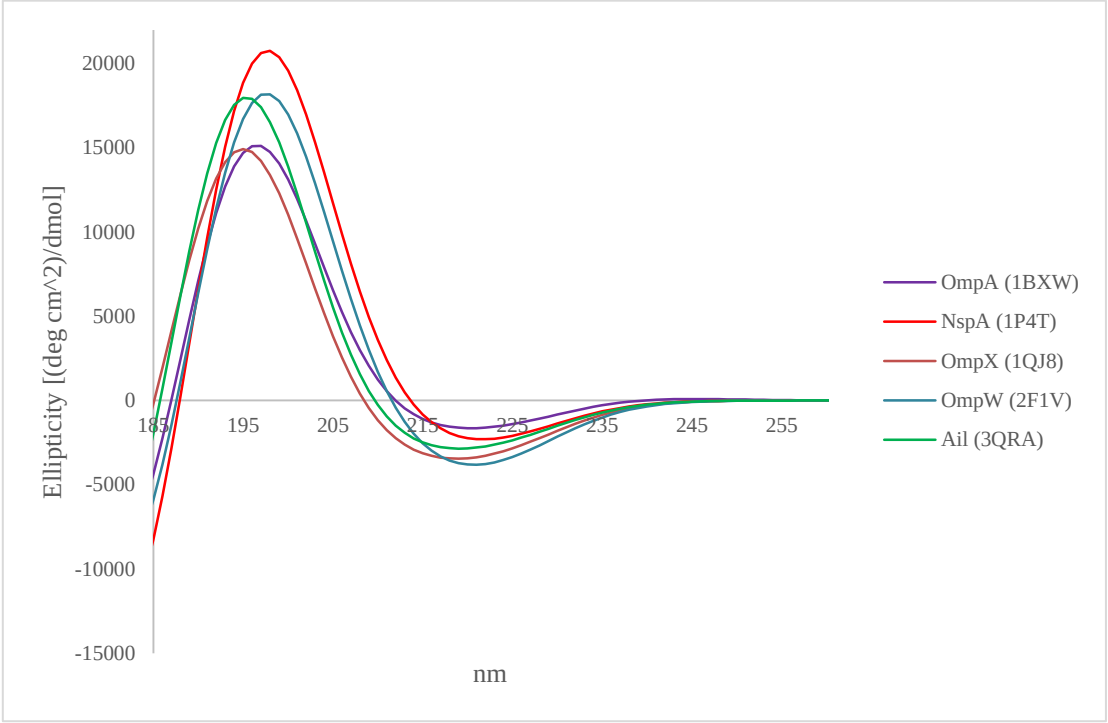


Figure 3.26. Predicted CD spectra of structurally solved 8-stranded β-barrels. CD spectra were generated using pdb files and the online server Dichrocalc.

3.10. Small angle X-ray scattering of *B. burgdorferi s.l.* OmpA-like proteins

3.10.1 Introduction to small angle X-ray scattering

Small-angle scattering of X-rays (SAXS) and neutrons (SANS) are well-established techniques commonly used to elucidate structural information from a wide variety of samples, including nanoparticles, emulsions, metal alloys, synthetic polymers and importantly biological macromolecules in solution (Svergun, 2007). First applications of the techniques began in the late 1930's following the early studies of metallic alloys by Guinier (Guinier, 1939). This small angle scattering of X-rays close to the primary beam provides a unique scattering profile and when this profile is acquired from a monodisperse solution of non-interacting particles the data carries important information about size and shape. By the 1960's the method was becoming increasingly important for the study of biological macromolecules in solution as it bypassed the requirement of crystalline material yet yielded low resolution structural information of around 1-2 nm and provided insight of the overall shape and internal structure (Svergun, 2007). By the 1980's this interest decreased as progress in other structural tools furthered. The challenging problem with the small angle scattering of biological molecules was how one may take a one-dimensional scattering pattern and then extract three-dimensional structural information; the experiments were therefore limited to simple measurements such as volume and calculation of the radius of gyration. A breakthrough during the 1990's however changed these limitations, data analysis methods vastly improved and alongside further advances in instrumentation *ab initio* shape determination using rigid body refinement became possible.

3.10.2. SAXS experimental set-up

X-rays are part of the electro-magnetic spectrum alike that of visible light. However, the wavelength is much shorter. As an alternating electric field causes an alternating magnetic field the waves propagate at right angles from each other. The interactions of these X-rays with matter, results in both absorption and scattering, a fraction may pass straight through the sample, some may be absorbed leading to other energy releases and a further portion will be scattered into alternative directions of propagation (Svergun, 2007). For structural experiments hard X-rays with wavelengths (λ) of 0.10-0.15 nm and energies of around 10keV are used.

A typical biological SAXS experiment utilizes a monochromatic and collimated X-ray beam which illuminates macromolecules in solution and their scattering is recorded by a detector. The detector is shielded from the intense incident beam by the beam stop and acts to prevent masking of the weak

scattering from the sample. As the macromolecules are in solution and not crystalline they are arranged in no particular orientation. Scattering is therefore normally presented as a radially averaged one dimensional curve following suitable background subtraction. This curve is often described by the intensity (I) as a function of the magnitude of q of the scattering vector as shown in equation 5. This curve reveals several important parameters including information regarding size, overall shape and oligomeric state of the molecule (Jacques and Trewhella, 2010).

Statistically the quality of the obtained scattering data for a given particle improves with increasing intensity. However, every detector will have a designated saturation value. Within an experimental set-up, once this saturation point is met any further counts may be lost in overflows. The detector saturation value for any detector often constitutes as the maximum exposure time for any given sample. Repeating short frames or exposures is often the chosen method to improve data quality rather than increasing the exposure times, although this varies for each detector type and experimental set-up.

Contrast is equally important in SAXS experiments, should the contrast between the target particle and the solvent is zero, scattering is negligible and increasing exposure times or repeating frames will not improve visibility. Most experiments are designed so that the concentrations of the target molecules are far higher than the constituents of the solvent. However, contrast is particularly important when working with solvents containing detergents and particle samples at low concentrations.

For accurate modelling using small angle scattering data the sample must be monodisperse, as the scattering generated from proteins is directly related to the square of their scattering length (Svergun et al., book) and aggregates can easily contribute disproportionately to the signal. Aggregation is easily identified by evaluation of the scattering profile at low q-values and manifests itself as a short upturn or increase in scattering. For monodisperse systems the background reduced intensity is directly proportional to the scattering from one molecule which has been averaged over all orientations (Svergun, 2007). The extent of the target signal can be described as directly relative to the number of particles within the sample that are illuminated by the X-ray beam (Petoukhov and Svergun, 2007).

$$Q = \frac{4\pi \sin(\theta)}{\lambda}$$

Equation 5. SAXS pattern representation as scattered intensity as a function of the magnitude of the scattering vector. Where θ is the angle between the X-ray bean and the detector and λ is the wavelength of the X-rays.

In general most small angle scattering experiments are carried out between 1-10 mg/ml. However, a range of concentrations are often used in order to determine the presence of interparticle interactions where in contrast to aggregation they are identified by a sharp downturn a low q.

3.10.3. Data analysis and Ab initio methods

Detected scattering profiles are normally subject to a radial integration, producing a scattered intensity curve. Following appropriate solvent reduction these scattering curves can be used to determine a number of parameters including, the radius of gyration (*Rg*), molecular mass (MM), the maximum diameter of the particle (Dmax), the extrapolated intensity at zero scattering angle I(0) and the overall shape or a low resolution molecular envelope.

Prior to modelling the scattering data is subjected to a Fourier transform normally using the program GNOM (Svergun, 1992), this transform takes the one-dimensional scattering curves and can calculate the particle distance distribution function p(r) and prepares output files for modelling. Following the transform the output files can be used to determine a low resolution shape to the randomly orientated particles within the solution. Popular modelling programs include, DAMMIN (Svergun, 1999), DAMMIF (Franke & Svergun, 2009) and GASBOR (Svergun et al., 2001). Both DAMMIN (Svergun, 1999) and DAMMIF (Franke & Svergun, 2009) begin by creating a search volume which represents the particle in question, this search volume, which is often a sphere has a radius of gyration larger than what is expected of the particle and is filled with dummy atoms. The program assigns each dummy atom as either the solute (protein) or the solvent and from taking an arbitrary start model the program uses simulated annealing in attempt to produce a compact, interconnected model which would generate a scattering pattern in close agreement to that of the experimental data. These initial modelling programs are normally carried out 10-20 times, each with an independent arbitrary initial model. This method results in the generation of a wide selection of models which all match the experimental scattering data.

Following the creation of the initial models, DAMAVER (Volkov & Svergun, 2003) is used to align, average and refine the envelopes. The DAMAVER (Volkov & Svergun, 2003) package is combined of a series of programs including, damsel, damsup, damaver and damfilt which are summarised in table 3.5. Following the averaging and refinement of the models by the DAMAVER (Volkov & Svergun, 2003) suite the final fit of the experimental data is often somewhere between the DAMAVER (Volkov & Svergun, 2003) and DAMFILT (Volkov & Svergun, 2003) envelope.

Program	Function	Reference
PRIMUS	An interface for performing manipulations of the scattering data including, subtraction, averaging, merging, extrapolation to zero concentration. Generation and evaluation of Guiner, Porod and Kratky plots.	Konarev et al., 2003
GNOM	An indirect transform program which takes one dimensional scattering curves and evaluates the P(r) and prepares .out files for modelling.	Svergun, 1992
DAMMIN	A tool for ab-initio modelling of SAXS data based on bead modelling and simulated annealing.	Svergun, 1999
DAMMIF	An improved ab-initio modelling program based upon the foundations of DAMMIN but implementing a new algorithm to speed up reconstruction by a factor of 25-40.	Franke & Svergun, 2009
SUPCOMB	A program used to superimpose one 3D structure onto another and calculates a normal spatial discrepancy (NSD) which can be used to determine the best alignments between models.	Kozin & Svergun, 2001
DAMSEL	The program takes all generated models in PDB format and superimposes all possible pairs using SUPCOMB then generates an output table which determines which models are outliers and selects a reference model. Models are considered as outliers if the NSD exceeds two standard deviations away from the calculated mean.	Volkov & Svergun, 2003
DAMSUP	A program that aligns a set of models excluding the outliers to a selected reference model and calculates standard deviation from the NSD.	Volkov & Svergun, 2003
DAMAVER	The program averages models that have been aligned by DAMSUP and computes a probability map.	Volkov & Svergun, 2003
DAMFILT	Can take the DAMAVER file and filter the model by removal of low occupancy regions or loosely connected atoms, this model doesn't always reflect the true size of the experimental data.	Volkov & Svergun, 2003

Table 3.5. Programs used for SAXS data analysis. A summary of the popular programmes used for the manipulation and analysis of scattering data.

3.10.4. Methodology control protein – Lysozyme

3.10.4.1 Background subtraction of lysozyme

Other than poor sample preparation the most common source of error during SAXS experiments is inappropriate background subtraction (Jacques and Trewhella, 2009). *Ab initio* modelling methods are particularly sensitive to such errors and can often deliver unexpected reconstructions when given poor data to build from. With this in mind lysozyme and ribonuclease A were employed as control proteins, they were selected due to their availability and known crystal structure. The control experiments were used to ensure correct background subtractions were being made and to investigate any possible changes to the constructed envelopes upon the addition of the detergent LDAO. Buffer blank samples of 0.3M NaCl, 50mM tris base, pH8 both with and without 0.1% v/v LDAO were subjected to 10 exposure frames each lasting 2400 seconds. The scattering profiles were independently recorded and integrated using the inbuilt Bruker software package. Lysozyme purchased from Sigma was reconstituted in the same buffer as described above, made up at 7 mg/ml both with and without 0.1% v/v LDAO. Lysozyme samples were exposed to X-rays in the same manner as the above buffer blanks and the scattering integrated by chi. Background subtraction of lysozyme with and without detergent was calculated using Primus (Konarev et al., 2003), by the subtraction of the scattering of the corresponding buffer blank which are shown in figure 3.27.

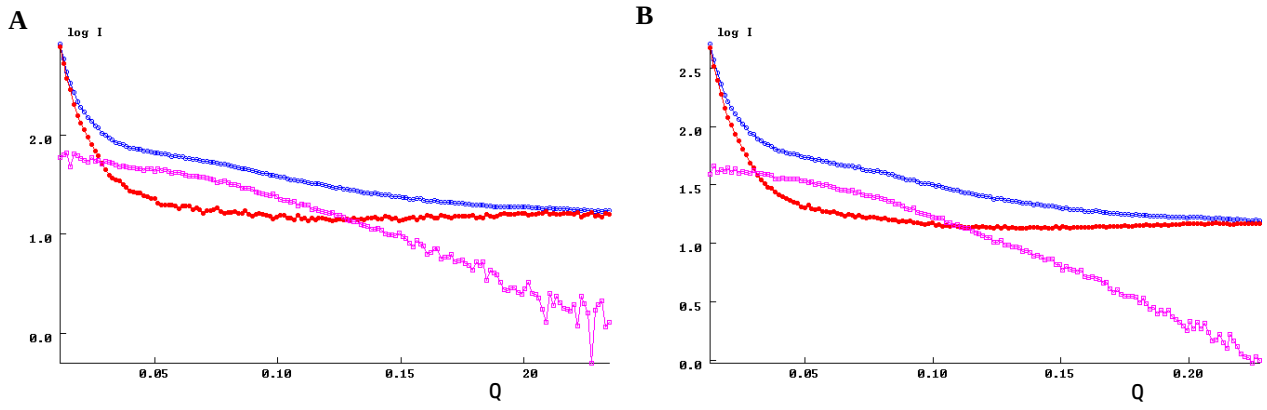


Figure 3.27 – Small angle scattering curves produced by 7 mg/ml lysozyme with and without 0.1% v/v **LDAO**. Recorded scattering data displayed as log of intensity over the scattering vector (Å). A) Raw scattering intensity plot of lysozyme and blank without LDAO. Red line is the scattering recorded for the blank sample (0.3M NaCl, 50mM tris base, pH8). Blue line shows the raw scattering recorded for lysozyme at 7 mg/ml in 0.3M NaCl, 50mM tris base, pH8. Purple line represents the adjusted scattering of lysozyme following buffer blank background subtraction. B) Raw scattering intensity plot of lysozyme and blank with the detergent LDAO. Red line is the scattering recorded for the blank sample (0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8). Blue line shows the raw scattering recorded for lysozyme at 7 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. Purple line represents the adjusted scattering of lysozyme following buffer blank background subtraction.

3.10.4.2. Calculating the radius of gyration (*Rg*) of lysozyme

Following the collection of small angle scattering data it is possible to perform many types of analysis but within many SAXS experiments the initial calculation of a particles radius of gyration is particularly useful. The radius of gyration provides information regarding to the mass distribution or the overall spread of the molecule and can be defined as the root mean square distance of the collection of atoms from their shared centre of gravity. As this parameter is based upon the spread of a particles constituents, molecules which have the same volume but hold different shapes will have different radius of gyration values (Jacques and Trewhella, 2009). For most spherical molecules the radius of gyration can be calculated by a Guinier approximation which is described in equation 6. The approximation is based upon the slope of a Guinier plot at low q and allows for the estimation of two parameters, which include the radius of gyration and the extrapolated intensity at zero scattering angle, I(0).

From the Guinier plot important information about the state of the protein within the solution can be extracted. Firstly any upturn at low q^2 that cannot be explained by the beam stop is often a sign of aggregation. If this upturn is so large as to affect the overall slope the data must be discarded and no further analysis made, as the influence of this type of aggregation can greatly affect the scattering pattern and is extremely difficult to remove. Downturns within this region are normally indicative of interparticle interference which can often be alleviated by reducing the protein concentration.

$$R_g = \sqrt{-3m}$$

Equation 6. Calculating the radius of gyration of a spherical object by Guinier analysis. Where m = the slope of the line.

In practical terms the Guinier approximation is often applied using a function in Primus (Konarev et al., 2003) known as Auto*Rg* or alternatively is manually fitted by the user who selects the linear region within the slope of the Guinier plot. The Auto*Rg* function within Primus (Konarev et al., 2003) calculates the radius of gyration by selecting a data range which is suitable for the application of the Guinier approximation. It begins by searching the initial portion of the input data and discards

anything that presents an upward or downward trend. Data is selected after these regions within the range where the scattering intensity begins to decay by an order of magnitude. The program them draws a cubic parabola within the chosen range in log scale of intensity and searches for curvature which could suggest a non-monodisperse system. This range is continuously refined until a suitable data range has been selected for Guinier analysis. Following the removal of any upturns and downturns and assessment of the monodispersity of the sample the program begins to select possible Guinier intervals which are normally between 5 to 15 data points. For each interval the program scores the linear fit whilst working within the limits of smin*Rg*> 1 and smax*Rg*< 1.3. Once the best intervals are found the radius of gyration is calculated and an assessment made on the accuracy of the calculation based upon the quality of the intervals used. The Auto*Rg* function can also give an idea of the overall quality of the data it does this by scoring several criteria which include,

a. How many consistent intervals were found
b. Where the s*Rg* conditions weakened by selection
c. How many initial points were discarded
d. Was there indication of effects like aggregation or interparticle interference
e. How accurate was the calculated radius of gyration.

In most cases assessment of the radius of gyration using the auto function is a speedy and highly reliable tool. However, there are some circumstances where the program may fail, these may include experiments where the molecule is very large or not globular and the default S*Rg* limits of 1-1.3 are not suitable. In some cases the upper S*Rg* limit can be increased to 1.5 and on some occasions even higher. Auto*Rg* may also struggle with scattering data containing upturns due to the beamstop, in these cases it often calculates a reasonable radius of gyration but assesses the quality as low.

Alternatively Primus (Konarev et al., 2003) offers the function to display the scattering data as a Guinier plot whereby the user can manually select the linear region within given S*Rg* limits allowing the calculation of the radius of gyration. The Guinier analysis for the control protein lysozyme is shown in figure 3.28. When manually selecting the Guinier region the goal is to remove any nonlinear points beginning in the low-q region in order to select a Guinier region within the desired S*Rg* limits. Working within the s*Rg* limits is normally crucial to ensure the estimated radius of gyration and I(0) are within 10% of the true value. As a Guinier analysis is an approximation of the SAXS curve the determined parameters are prone to normal estimation errors. Also any curvature in the residuals

shown in figure 3.28 is representative of non-linear behaviour and the possibility of a non-monodisperse system.

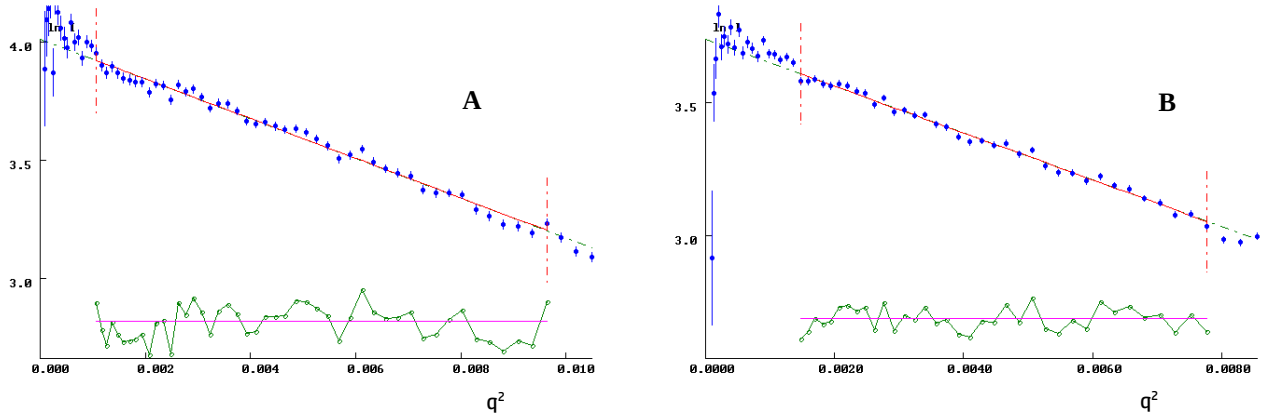


Figure 3.28. Manual Guinier plots of the control protein lysozyme generated using Primus. The scattering data is plotted as In[I(q)] vs q^2. The blue data points represent the scattering data, the red line shows the Guinier fit and the green data shows the corresponding residuals. A) Guinier plot analysis of the scattering data from lysozyme without detergent. B) Guinier analysis of lysozyme in 0.1% v/v LDAO.

For the first control protein lysozyme both Auto*Rg* and manual Guinier analysis was applied to both of the data sets and the calculated radius of gyration values are summarised in table 3.6. Interestingly the radius of gyration is reasonably consistent regardless of the method of calculation but most importantly it does not substantially change with the addition of LDAO which is a major concern when using detergents in SAXS experiments. The calculated radius of gyration of lysozyme which is most accurately calculated using GNOM (Svergun, 1992) was deemed to be 16.09 Angstrom which correlates particularly well with literature data of 16.2 at 20^{o}C under the conditions tested (Arai & Hirai 1999).

Parameter Calculated	***Rg* lysozyme without LDAO**	***Rg* lysozyme in LDAO**
Primus Auto *Rg*	15.22 +-0.08 (s*Rg* 0.52-1.256)	16.34 +-0.31 (s*Rg* 0.674-1.302)
Manual Guinier *Rg*	15.90 +-0.104 (s*Rg* 0.519-1.56)	16.3 +-0.101 (s*Rg* 0.655-1.490)
Reciprocal space *Rg*	16.09	16.08
Real space *Rg*	16.10	16.09

Table 3.6. A summary of the calculated *Rg* for lysozyme with and without the detergent LDAO. On the whole the calculated *Rg* for lysozyme appears to correspond well regardless of the method used to determine the parameter. The value calculated using Auto*Rg* appears to diverge a from other methods possibly due to issues with the selection of the Guinier region.

3.10.4.3. Kratky analysis of lysozyme

A Kratky analysis of scattering data can give a wealth of information about the 'unfolded-ness' of a sample and is often used to identify disordered states from globular particles, examples are shown in figure 3.29A. As globular macromolecules follow Porod's law they display a bell-shaped curve. In contrast any extended molecules or unfolded proteins lack a defined maxima and often plateau or increase within the high q range (Putnam et al., 2007). The Kratky plot or $I(q)*q^2$ versus q was produced for the lysozyme scattering data using sasplot within the Primus (Konarev et al., 2003) software and is shown in figure 3.29B. The Kratky appears as a Gaussian-like bell shaped curve with a well-defined maximum, this suggests that the protein is folded under the current conditions tested and is likely to be a globular protein.

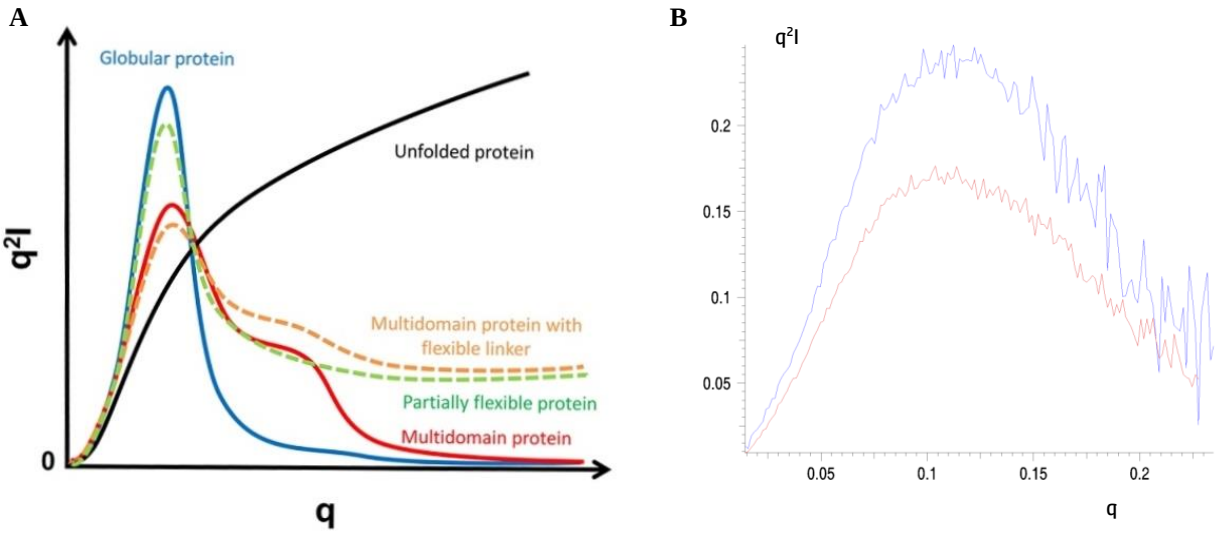


Figure 3.29. Kratky plot analysis. A) Schematic representation of possible Kratky plots adapted from literature issued by Standford Synchrotron Radiation Lightsource**. B)** Kratky plot of lysozyme. $I(q)*q^2$ plot of the lysozyme scattering data in buffers with and without LDAO. Blue – Lysozyme in 0.3M NaCl, 50mM tris base, pH8. Red – Lysozyme in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. The Kratky plot of a Gaussian-like bell shape is indicative of a folded molecule.

3.10.4.4. GNOM analysis and evaluation of the pair-distance distribution function - P(r)

After the initial analysis of the integrated SAXS data, GNOM (Svergun, 1992) from the ATSAS (Svergun, 1992, Franke & Svergun, 2009, Volkov & Svergun, 2003, Kozin & Svergun, 2001, Svergun, 1999) software package is often utilised as an indirect transform program which takes one-dimensional scattering curves and calculates a distance distribution function or P(r) it also prepares files for ab initio modelling. The P(r) function is a description applied to the paired set of all distances between points within an object. Within SAXS experiments this can be applied to describe the paired set of distances between all of the atoms within the macromolecular structure and is extremely useful in understanding conformational changes within a macromolecule as a small change made, for example by the binding of a ligand will result in a detectable change in the P(r) distribution. These small changes can cause alterations in the mass distribution of the macromolecule around its centre of gravity and it's radius of gyration and as this can be obtained from the P(r) function, comparisons of the radius of gyration of two different samples can expose possible conformational changes within the macromolecule.

GNOM (Svergun, 1992) was used to action a Fourier transform of the lysozyme background subtracted scattering data, a suitable Rmax of 47 was chosen, this selection was based upon the smallest value possible that gave the data or curve a smooth return to zero. It was observed that any further decrease to the Rmax produced a curve which plummeted to zero and any increase beyond 47 produced a stretched tail. In general terms the Rmax should be larger than the expected particle size and defines the bounds for the possible particle size. Globular macromolecules should display a P(r) function with a single peak while stretched or unfolded proteins have long tails at large r or in some cases display multiple peaks alike what is observed for multidomain proteins. Examples of possible P(r) functions are discussed further in figure 3.30.

The radius of gyration calculated using a Guinier approximation relies upon a small portion of the data from a SAXS experiment, in contrast the same parameter calculated from the P(r) distribution uses all of the experimental curve in real space and is often considered as the most accurate interpretation. Both lysozyme data sets with and without detergent were independently processed by GNOM (Svergun, 1992) and are shown in figure 3.31.

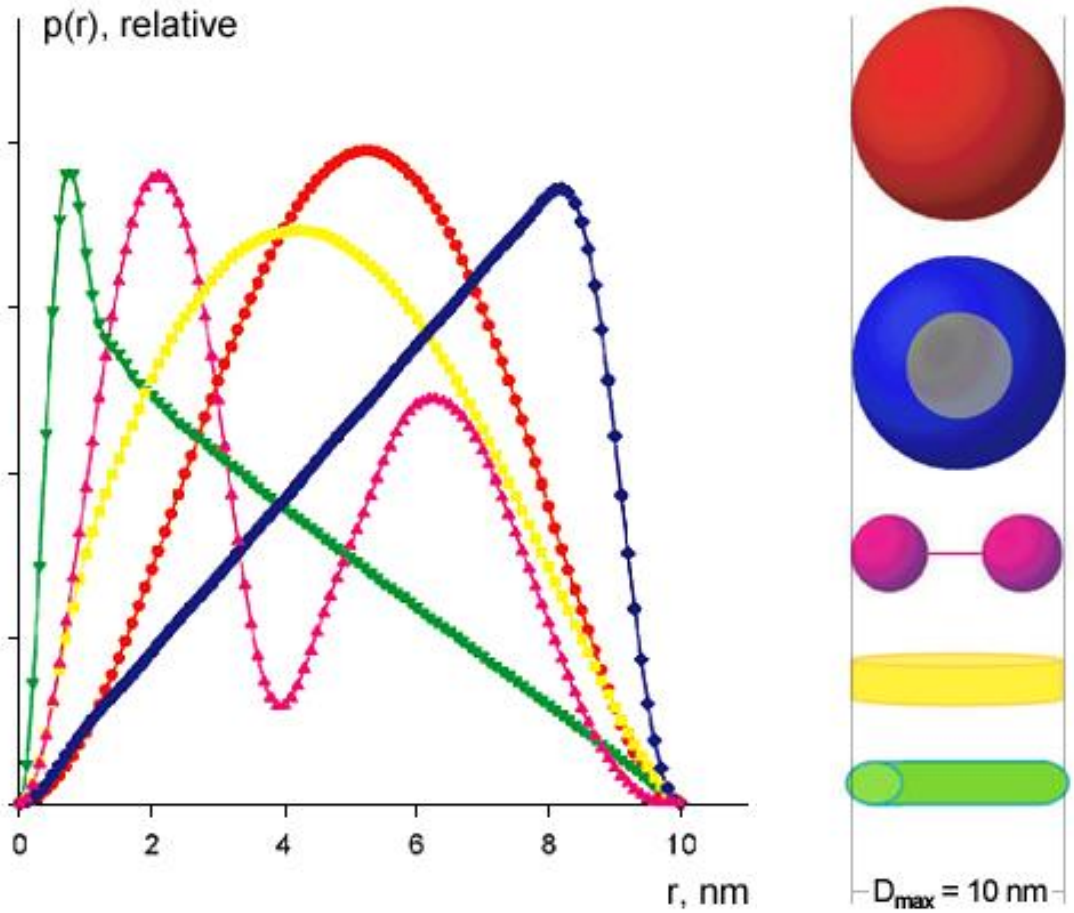


Figure 3.30. Examples of possible P(r) distribution plots. The relationship between relative P(r) and actual molecular shape. Adapted from SAS studies of biological macromolecules in solution by Svergun & Koch. (2003).

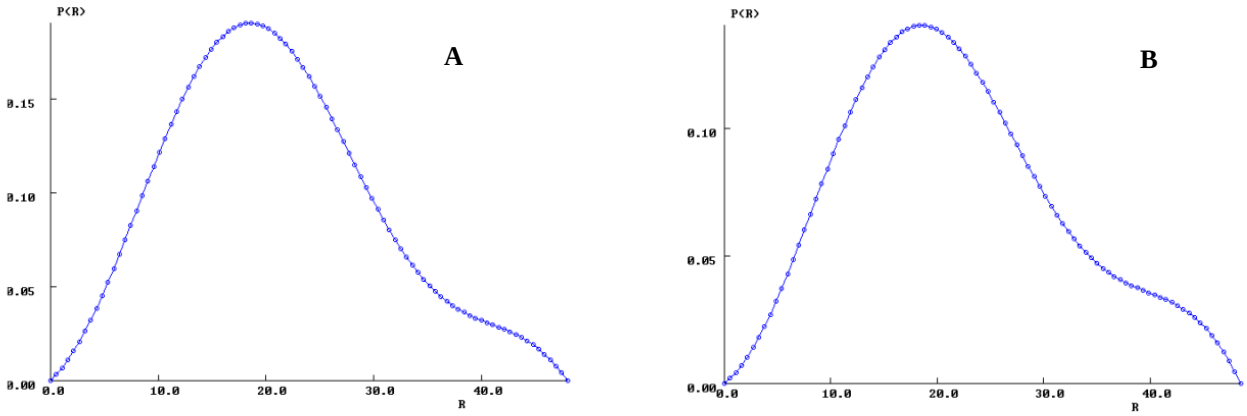


Figure 3.31. P(r) distribution plots for lysozyme with and without LDAO. Plots for both data sets were generated using GNOM (Svergun, 1992) using the full data set and an Rmax of 47. A) P(r) of lysozyme at 7 mg/ml in 0.3M NaCl, 50mM tris base, pH8. B) P(r) of lysozyme at 7 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8.

3.10.4.5. Generating the lysozyme molecular envelopes

Output files (.out) from GNOM (Svergun, 1992) were sent to DAMMIF (Franke & Svergun, 2009), an ab initio modelling program which was performed twenty times for each lysozyme data set, the models generated are shown in figures 3.32 and 3.33. DAMMIF (Franke & Svergun, 2009) is an improved version of a previous program, DAMMIN (Svergun, 1999). Both programs use a bead modelling method for rapid shape determination of small angle scattering data. The method uses a collection of densely packed beads to represent a particle within a search volume; each bead can be part of the macromolecule or part of the solvent. Using an arbitrary starting model the programs use simulated annealing to develop a compact model which yields a theoretical scattering pattern which in turn can be compared to the experimental data.

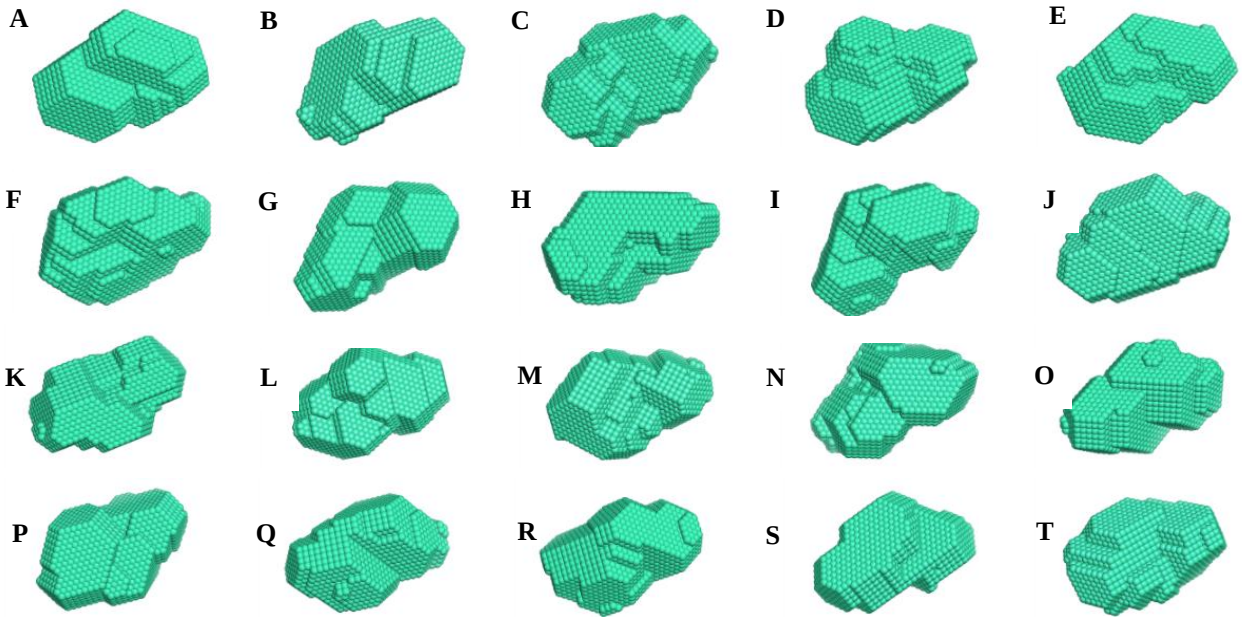


Figure 3.32. Molecular envelopes for lysozyme without LDAO. Data collected from a lysozyme sample at 7 mg/ml in 0.3M NaCl, 50mM tris base, pH8. Molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models.

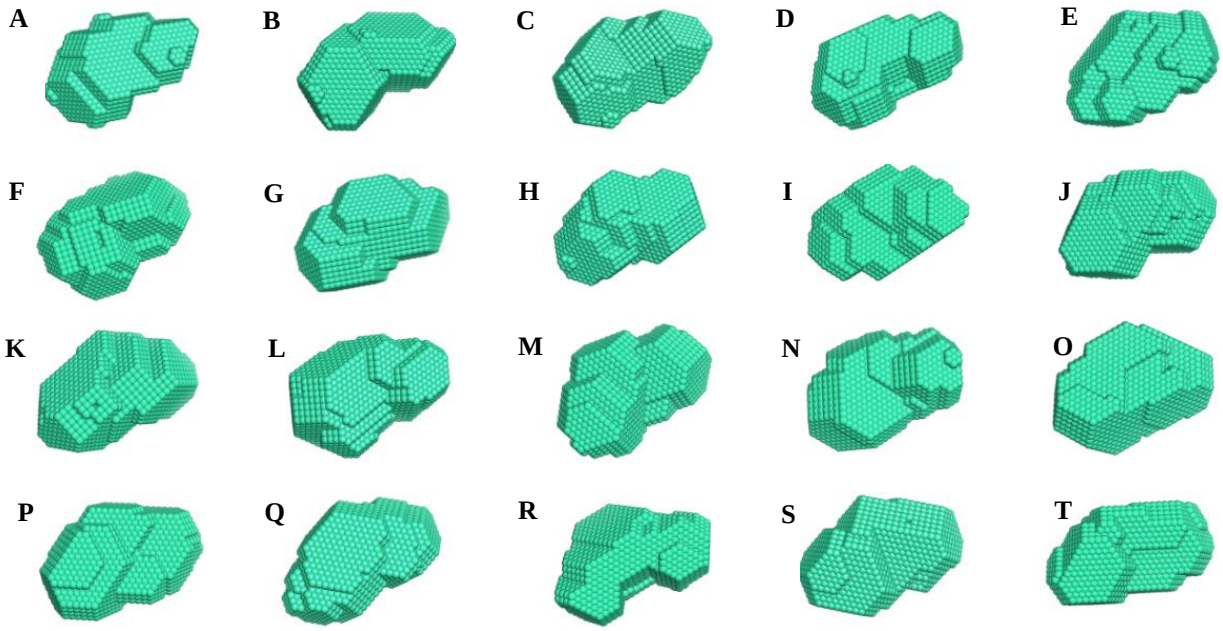


Figure 3.33. Molecular envelopes of lysozyme in 0.1% LDAO. Data collected from a lysozyme sample at 7 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. Molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models.

3.10.4.6. Evaluation and refinement of the lysozyme models

Assessment of the generated models was made using DAMAVER (Volkov & Svergun, 2003) and is summarised in table 3.7. DAMAVER (Volkov & Svergun, 2003) is a suite of programs used to align *ab initio* models allowing selection of the most probable solution from which, an averaged model is built. The group of programs begins with DAMSEL (Volkov & Svergun, 2003) which uses the SUPCOMB (Kozin & Svergun, 2001) method to superimpose the low resolution 3D structures and then begins to determine the most probable solution and eliminate outliers.

DAMSUP (Volkov & Svergun, 2003) then takes all the models excluding the outliers and aligns the envelopes with the model that was selected as most probable. DAMAVER (Volkov & Svergun, 2003) then averages these aligned models and computes a probability map which is then filtered by DAMFILT (Volkov & Svergun, 2003) at a given cut off volume. Following removal of outliers and the superimposition of all models upon the reference envelope DAMAVER (Volkov & Svergun, 2003) produced an averaged probability map which was filtered by the removal of low occupancy regions by DAMFILT (Volkov & Svergun, 2003). Both DAMFILT (Volkov & Svergun, 2003) and DAMAVER (Volkov & Svergun, 2003) models were used to produce the final envelope with a cartoon of the lysozyme crystal structure (1LYZ – Diamond, 1974) docked into the bead model. PDB coordinates for lysozyme were manually docked into the molecular envelope using Pymol (Schrödinger, 2010). The final refined envelopes are shown in figure 3.34 and 3.35.

Solution	DAMSEL NSD average (Standard deviation of NSD = 0.020)	Include/Discard?	DAMSUP NSD superimposition value	Indicative fit (Chi squared)
Lysozyme (No LDAO)				
A	0.539	**Reference**	0.000	1.960
B	0.593	**Include**	0.572	1.962
C	0.566	**Include**	0.571	1.961
D	0.560	**Include**	0.535	1.960
E	0.552	**Include**	0.457	1.959
F	0.571	**Include**	0.530	1.961
G	0.610	**Discard**	-----	1.961
H	0.544	**Include**	0.524	1.963
I	0.617	**Discard**	-----	1.960
J	0.560	**Include**	0.521	1.960
K	0.553	**Include**	0.542	1.959
L	0.557	**Include**	0.502	1.958
M	0.571	**Include**	0.554	1.961
N	0.565	**Include**	0.548	1.962
O	0.566	**Include**	0.554	1.962
P	0.566	**Include**	0.554	1.959
Q	0.565	**Include**	0.495	1.959
R	0.562	**Include**	0.554	1.959
S	0.555	**Include**	0.491	1.960
T	0.552	**Include**	0.530	1.958
Lysozyme (0.1% v/v LDAO)				
A	0.551	**Include**	0.546	1.578
B	0.576	**Include**	0.538	1.588
C	0.553	**Include**	0.505	1.575
D	0.548	**Include**	0.532	1.578
E	0.542	**Include**	0.517	1.575
F	0.558	**Include**	0.525	1.577
G	0.522	**Reference**	0.000	1.577
H	0.533	**Include**	0.472	1.576
I	0.545	**Include**	0.515	1.578
J	0.528	**Include**	0.506	1.576
K	0.532	**Include**	0.516	1.576
L	0.546	**Include**	0.521	1.581
M	0.536	**Include**	0.513	1.581
N	0.538	**Include**	0.526	1.572
O	0.562	**Include**	0.537	1.576
P	0.535	**Include**	0.493	1.575
Q	0.532	**Include**	0.516	1.576
R	0.610	**Discard**	-----	1.576
S	0.539	**Include**	0.509	1.582
T	0.539	**Include**	0.504	1.581

Table 3.7. A summary of the evaluation and refinement of lysozyme models by DAMSEL and DAMSUP. DAMSEL (Volkov & Svergun, 2003) determined the mean NSD as 0.556 and 0.546 for models generated with and without LDAO respectively. Envelopes G and I generated for lysozyme without LDAO and envelope R with LDAO were calculated as outliers and the data discarded. Data for envelope A and G again with and without LDAO were identified as the most probably solutions in each case and this data used as the reference sets for superimposition. DAMSUP (Volkov & Svergun, 2003) NSD values were calculated relative to the reference set. The chi squared value represents the indicative fit calculated by DAMMIF (Volkov & Svergun, 2003) for each generated model.

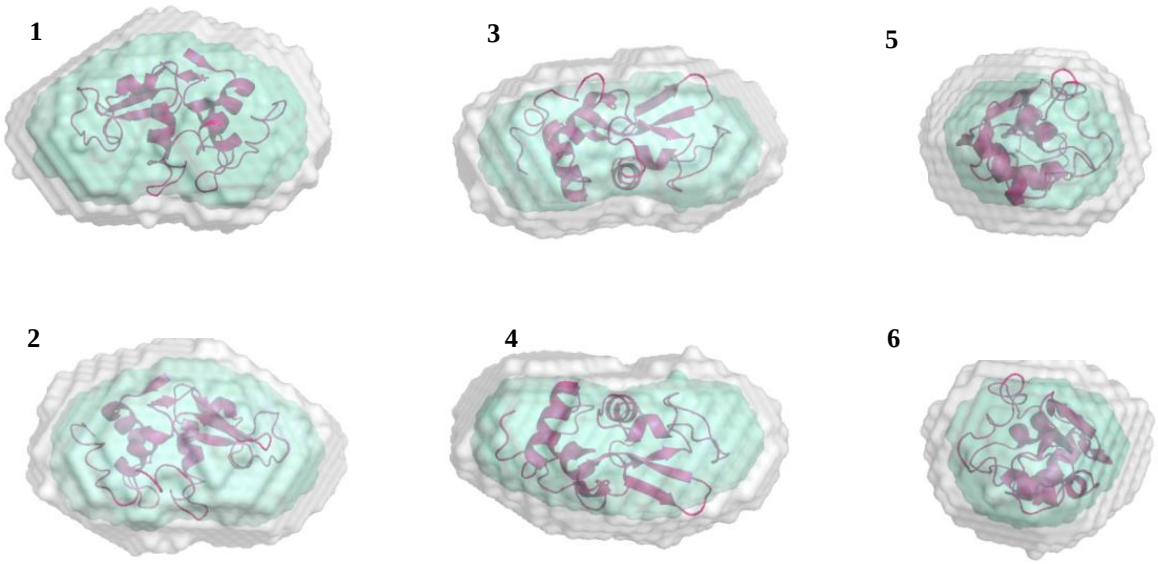


Figure 3.34. Refined and filtered molecular envelopes for lysozyme without LDAO. Final molecular envelopes for lysozyme at 7 mg/ml in 0.3M NaCl, 50mM tris base, pH8. The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope, the grey as the computed probability map generated by DAMAVER (Volkov & Svergun, 2003) and the pink ribbon cartoon is the known crystal structure of lysozyme (1LYZ) taken from the protein data bank. 1-2. Front and rear view. 3-4. Side view. 5-6. End on views of the molecule.

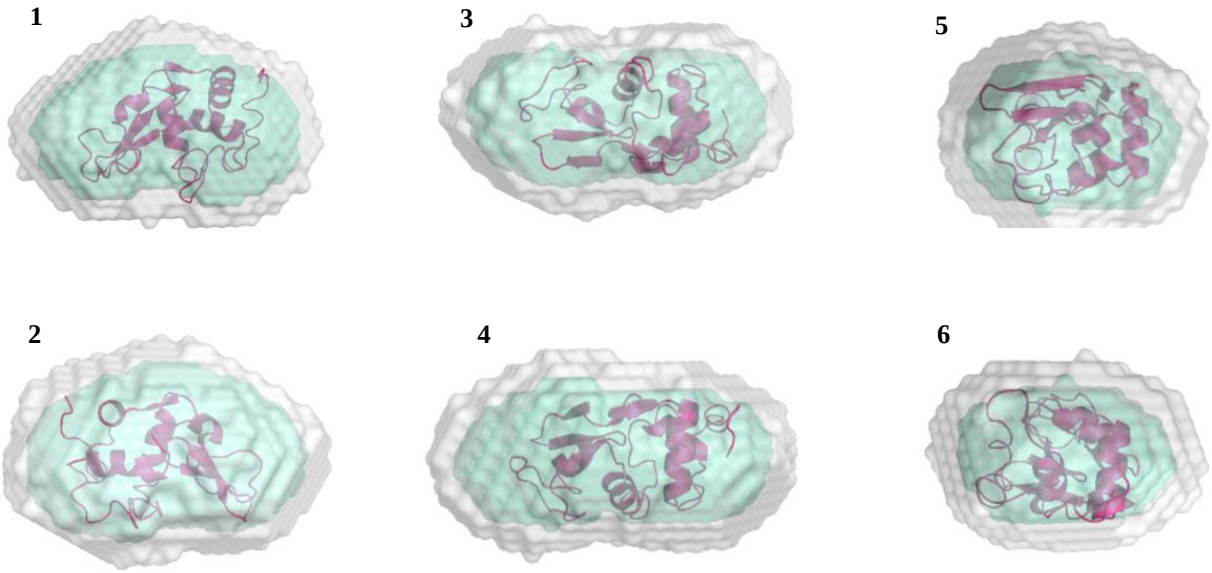


Figure 3.35. Refined and filtered molecular envelopes for lysozyme in 0.1% v/v LDAO. Final molecular envelopes for lysozyme at 7 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope, the grey as the computed probability map generated by DAMAVER (Volkov & Svergun, 2003) and the pink ribbon cartoon is the known crystal structure of lysozyme (1LYZ) taken from the protein data bank. 1-2. Front and rear view. 3-4. Side view. 5-6. End on views of the molecule.

3.10.4.7. Effects of LDAO upon the lysozyme molecular envelope

In order to deduce the effects of LDAO scattering upon the lysozyme models a final envelope for lysozyme in LDAO and without detergent was created. These DAMFILT (Volkov & Svergun, 2003) envelopes are shown in figure 3.36. The envelopes were overlaid alongside the mesh structure from published X-ray diffraction experiments. The envelopes correlate well and no major changes to the model can be identified when LDAO was added to the sample.

The data analysis of the control protein lysozyme was deemed suitable, the *Rg* of 16Å was not seen to change between the two data sets with and without the detergent LDAO and is very close to literature values which are summarised in table 3.8. Most importantly it was determined that the scattering caused by the detergent LDAO could be compensated for by the acquisition of good quality data used for background subtraction. The control experiments identified the need for extra care when preparing blanks with a dialysis matched blank being the most desired option.

Molecule	**Radius of gyration (Å)**
Lysozyme in 0.3M NaCl, 50mM tris base, pH8	16.085
Lysozyme in 0.1% v/v LDAO, 0.3M NaCl, 50mM tris base, pH8	16.095
Lysozyme literature data Arai & Hirai 1999	16.3

Table 3.8. A summary of the radius of gyration of lysozyme. Calculated *Rg* of lysozyme both with and without 0.1% v/v LDAO alongside the literature data. The calculated *Rg* with and without detergent are extremely close and both are relative to published literature values.

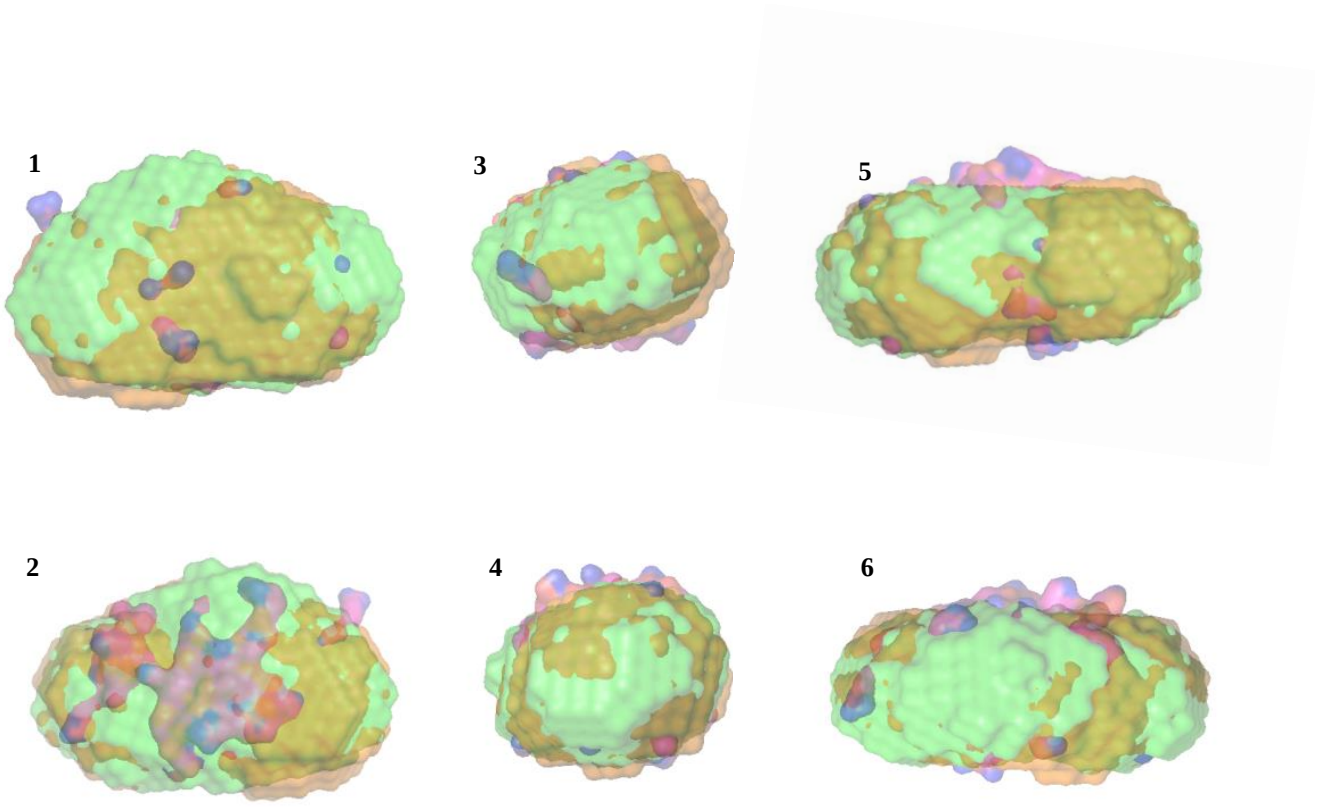


Figure 3.36. Assessment of the final molecular envelope of lysozyme by overlaying the models. Overlaid models of lysozyme generated in Pymol (Schrödinger, 2010) and showing the molecular envelope of lysozyme with LDAO (orange) and without (green), plus a surface representation of the lysozyme crystal structure (purple) taken from the protein data bank (1LYZ). 1-2. Front and rear view. 3-4. End on views. 5-6. Side view of the overlaid envelopes.

3.10.5. Methodology control protein – Ribonuclease A

3.10.5.1. Background subtraction of Ribonuclease A

Scattering experiments were carried out for bovine pancreatic ribonuclease A as a second control in order to monitor the effects of LDAO upon the determined molecular envelope. Ribonuclease A was purchased from Sigma Aldrich and reconstituted at 5 mg/ml in 0.3M NaCl, 50mM tris base at pH8 both with and without 0.1% v/v LDAO. The sample was subsequently centrifuged at 15,000 x g for 30 minutes before scattering data was collected at the B21 beamline, Diamond Light Source, Oxford. A total of 18 frames per sample were collected each lasting 3 seconds with the corresponding buffer blank sample collected both prior and after each sample. Following data integration the scattering generated for the blank buffer was subtracted from the scattering produced for the control sample, ribonuclease A. Subtractions were carried out as previously described for the lysozyme control using Primus (Konarev et al., 2003) and the resulting curves are shown in figure 3.37.

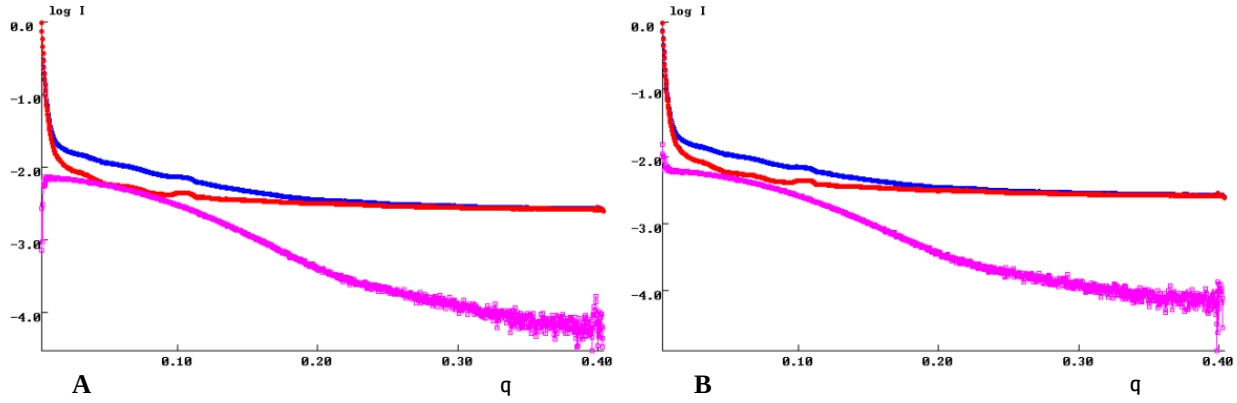


Figure 3.37. Small angle scattering curves produced by 5 mg/ml ribonuclease A with and without 0.1% LDAO. Recorded scattering data is displayed as log of intensity over the scattering vector q (Å). **A)** Raw scattering intensity plot of ribonuclease A and blank without LDAO in the samples. The red line is the scattering recorded for the blank sample (0.3M NaCl, 50mM tris base, pH8). The blue line shows the raw scattering recorded for ribonuclease at 5 mg/ml in 0.3M NaCl, 50mM tris base, pH8. The purple line represents the adjusted scattering of ribonuclease A following buffer blank background subtraction. **B)** Raw scattering intensity plot of ribonuclease A and blank with the detergent LDAO. The red line is the scattering recorded for the blank sample (0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8). The blue line shows the raw scattering recorded for ribonuclease A at 5 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8 and the purple line represents the adjusted scattering of ribonuclease A following buffer blank background subtraction.

3.10.5.2. Calculating the radius of gyration (*Rg*) of ribonuclease A

The radius of gyration was calculated both manually using a Guinier approximation and by using the Auto*Rg* function within the Primus software (Konarev et al., 2003). Guinier plots for ribonuclease A, both with and without detergent are shown in figure 3.38 and the calculated radius of gyration was a larger than the literature data of 14.3Å (Wang et al., 2008). From the Guinier analysis the sample was found to be of satisfactory quality and the linear region was easily determined. The radius of gyration for ribonuclease A both with and without LDAO was exactly the same and was determined to be 16Å by Auto*Rg* and 16.4Å using a manual fitting, with full details of s*Rg* limits used summarised in table 3.9.

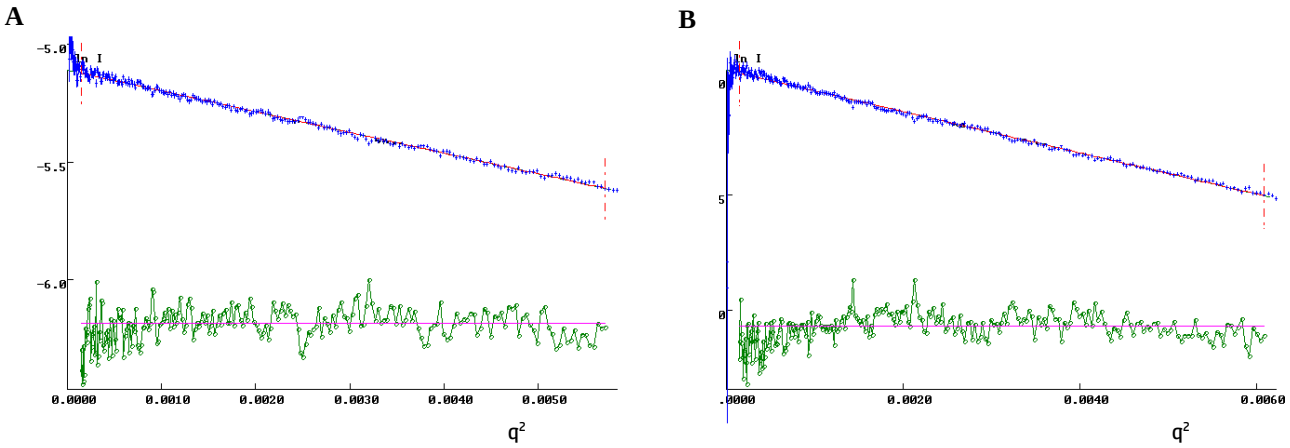


Figure 3.38. Manual Guinier plots of the control protein ribonuclease A generated using Primus. The scattering data is plotted as In[I(q)] vs q^2. The blue data points represent the scattering data, the red line shows the Guinier fit and the green data shows the corresponding residuals. A) Guinier plot analysis of the scattering data from ribonuclease A without detergent. B) Guinier analysis of ribonuclease A in 0.1% v/v LDAO.

Parameter Calculated	*Rg* ribonuclease A without LDAO	*Rg* ribonuclease A in LDAO
Primus Auto *Rg*	16.02 +-0.07 (s*Rg* 0.657-1.305)	16.04 +-0.05 (s*Rg* 0.502-1.305)
Manual Guinier *Rg*	16.4 +-2.95 (s*Rg* 0.196-1.3)	16.4 +-1.1 (s*Rg* 0.192-1.28)
Reciprocal space *Rg*	15.71	15.68
Real space *Rg*	15.71	15.68

Table 3.9. A summary of the radius of gyration for ribonuclease A with and without the detergent LDAO. The radius of gyration of ribonuclease A both with and without detergent was calculated by Guinier approximation and in real and reciprocal space. The values are in fair agreement but importantly there was no significant difference observed between the sample with LDAO.

3.10.5.3. Kratky analysis of ribonuclease A

A Kratky plot or $I(q)*q^2$ versus s was produced for both sets of ribonuclease A scattering data using sasplot within the Primus software (Konarev et al., 2003) which are shown in figure 3.39. The Kratky plot appears as a Gaussian-like bell shaped curve with a well-defined maximum point and the data returning towards zero which suggests a compact folded protein.

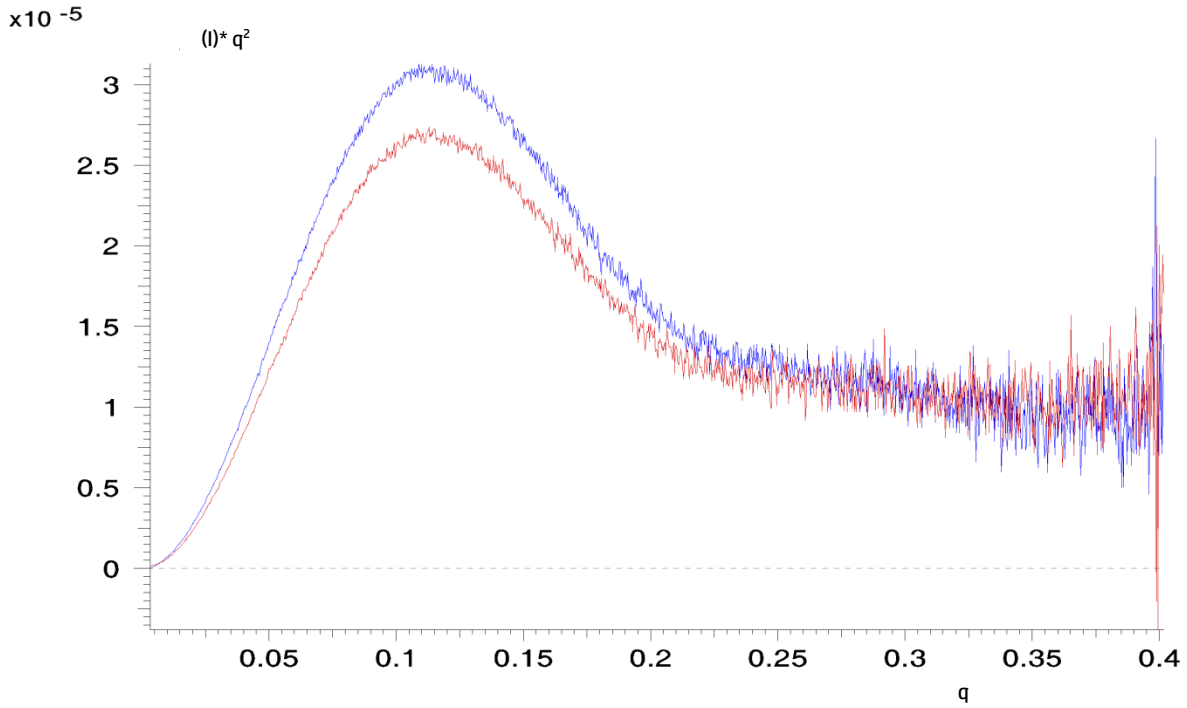


Figure 3.39. Kratky plots for ribonuclease A with and without 0.1% LDAO. $I(q)*q^2$ plot of the ribonuclease A scattering data in buffers with and without LDAO.

3.10.5.4. GNOM analysis and evaluation of the pair-distance distribution function P(r) of ribonuclease A

Following the initial analysis using Primus (Konarev et al., 2003) the full scattering curve was input into GNOM (Svergun, 1992) which actioned a Fourier transform of the ribonuclease A background subtracted scattering data. A suitable Rmax of 45 was chosen, this selection was based upon the smallest value possible that gave a smooth return to zero. Both ribonuclease A data sets were independently processed by GNOM (Svergun, 1992) and the P(r) functions are shown in figure 3.40. Both ribonuclease A data sets with and without detergent produced similar P(r) functions and were indicative of a compact protein.

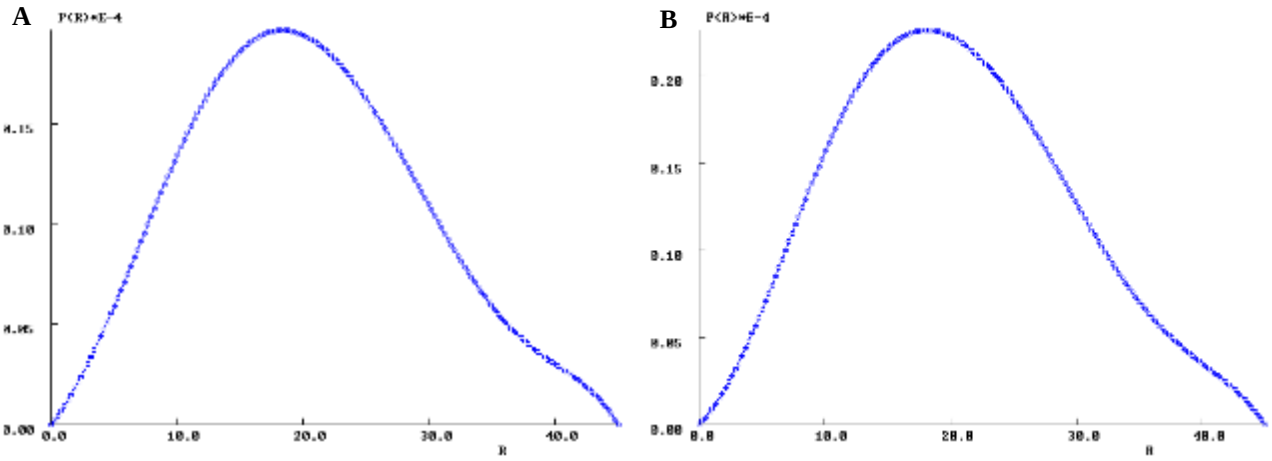


Figure 3.40. P(r) distribution plots for ribonuclease A with and without LDAO. Plots for both data sets were generated using GNOM (Svergun, 1992) using the full data set and an Rmax of 45Å. **A)** P(r) of Ribonuclease A at 5 mg/ml in 0.3M NaCl, 50mM tris base, pH8**. B)** P(r) of ribonuclease A at 5 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8.

3.10.5.5. Generating ribonuclease A molecular envelopes

Output files (.out) from GNOM (Svergun, 1992) were sent to DAMMIF (Franke & Svergun, 2009), an ab initio modelling program which was performed twenty times for each data set and the models generated are shown in figures 3.41 and 3.42. All modelling was performed as previously described for the first control protein lysozyme.

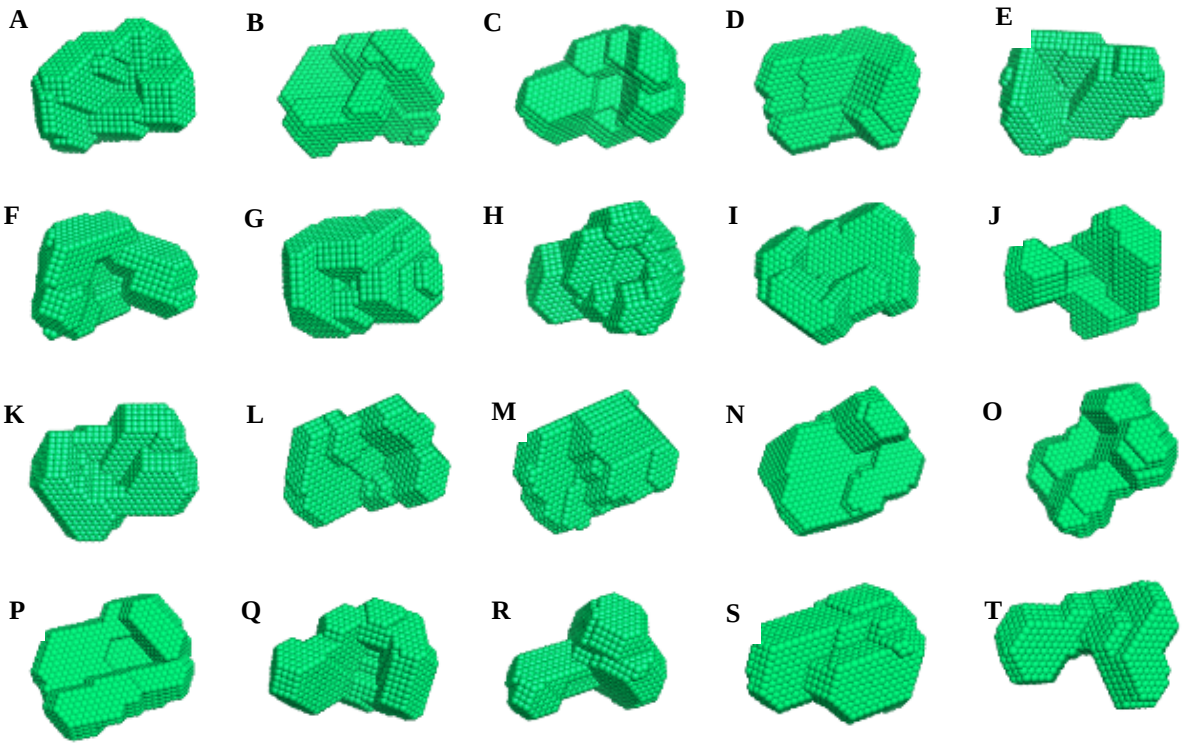


Figure 3.41. Molecular envelopes for ribonuclease A without LDAO. Data collected from a ribonuclease A sample at 5 mg/ml in 0.3M NaCl, 50mM tris base, pH8. Molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models.

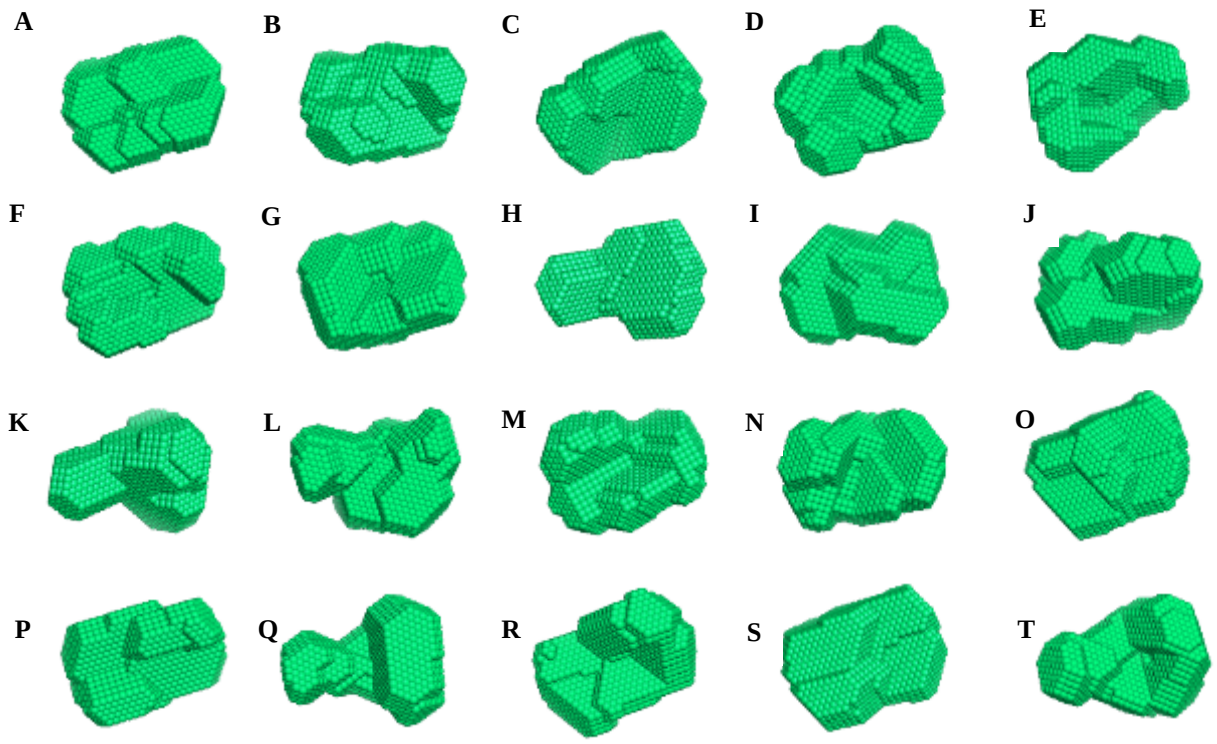


Figure 3.42. Molecular envelopes of ribonuclease A in 0.1% LDAO. Data collected from a ribonuclease A sample at 5 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. Molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models.

3.10.5.6. Evaluation and refinement of the ribonuclease A models

Assessment of the generated models was made using the DAMAVER suite of programs (Volkov & Svergun, 2003), the PDB files for the twenty generated models were submitted to DAMSEL (Volkov & Svergun, 2003). Following alignment of the models for ribonuclease A without LDAO it was determined that envelopes R and T should be discarded as their calculated NSD was beyond two points of standard deviation alike that of envelopes K and T for ribonuclease A envelopes with LDAO. Following removal of the outliers the remaining models were sent to DAMSUP (Volkov & Svergun, 2003) for further alignment and the NSD calculated a second time. All alignment data is shown in table 3.10.

Following alignment the remaining models were submitted to DAMAVER (Volkov & Svergun, 2003) where all models were averaged and a probability map computed, this averaged envelope was then taken by DAMFILT (Volkov & Svergun, 2003) for removal of low occupancy regions. The resulting final models are shown in figure 3.43.

Solution	DAMSEL NSD average (Standard deviation of NSD = 0.067 without LDAO and 0.106 with LDAO)	Include/Discard?	DAMSUP NSD superimposition value	Indicative fit (Chi squared)
Ribonuclease A (No LDAO)				
A	0.739	**Include**	0.592	4.829
B	0.853	**Include**	0.798	4.834
C	0.694	**Include**	0.646	4.783
D	0.736	**Include**	0.663	4.841
E	0.680	**Include**	0.642	4.804
F	0.737	**Include**	0.663	4.834
G	0.750	**Include**	0.589	4.803
H	0.756	**Include**	0.715	4.821
I	0.718	**Include**	0.513	4.809
J	0.803	**Include**	0.768	4.824
K	0.748	**Include**	0.632	4.822
L	0.673	**Reference**	0.000	4.813
M	0.690	**Include**	0.606	4.790
N	0.727	**Include**	0.619	4.789
O	0.739	**Include**	0.593	4.786
P	0.701	**Include**	0.564	4.786
Q	0.714	**Include**	0.667	4.821
R	0.916	**Discard**	-----	4.814
S	0.705	**Include**	0.675	4.825
T	0.890	**Discard**	-----	4.835
Ribonuclease A (0.1% LDAO)				
A	0.711	**Include**	0.535	3.339
B	0.726	**Include**	0.603	3.336
C	0.678	**Include**	0.506	3.327
D	0.736	**Include**	0.598	3.326
E	0.693	**Include**	0.615	3.335
F	0.683	**Include**	0.505	3.357
G	0.690	**Include**	0.528	3.326
H	0.836	**Include**	0.934	3.346
I	0.723	**Include**	0.563	3.312
J	0.767	**Include**	0.654	3.325
K	1.006	**Discard**	-----	3.339
L	0.774	**Include**	0.731	3.361
M	0.674	**Include**	0.536	3.346
N	0.683	**Include**	0.528	3.344
O	0.689	**Include**	0.547	3.322
P	0.709	**Include**	0.553	3.360
Q	0.839	**Include**	0.873	3.345
R	0.664	**Reference**	0.000	3.332
S	0.723	**Include**	0.550	3.324
T	1.052	**Discard**	-----	3.348

Table 3.10.A summary of the evaluation and refinement of ribonuclease A models by DAMSEL and DAMSUP. DAMSEL (Volkov & Svergun, 2003) determined the mean NSD as 0.748 and 0.753 for models generated with and without LDAO respectively. Envelopes R and T generated for ribonuclease A without LDAO and envelope K and T with LDAO were calculated as outliers and the data discarded. Data for envelope L and R again with and without LDAO were identified as the most probably solutions in each case and this data used as the reference sets for superimposition. DAMSUP (Volkov & Svergun, 2003) NSD values were calculated relative to the reference set. The chi squared value represents the indicative fit calculated by DAMMIF (Volkov & Svergun, 2003) for each generated model.

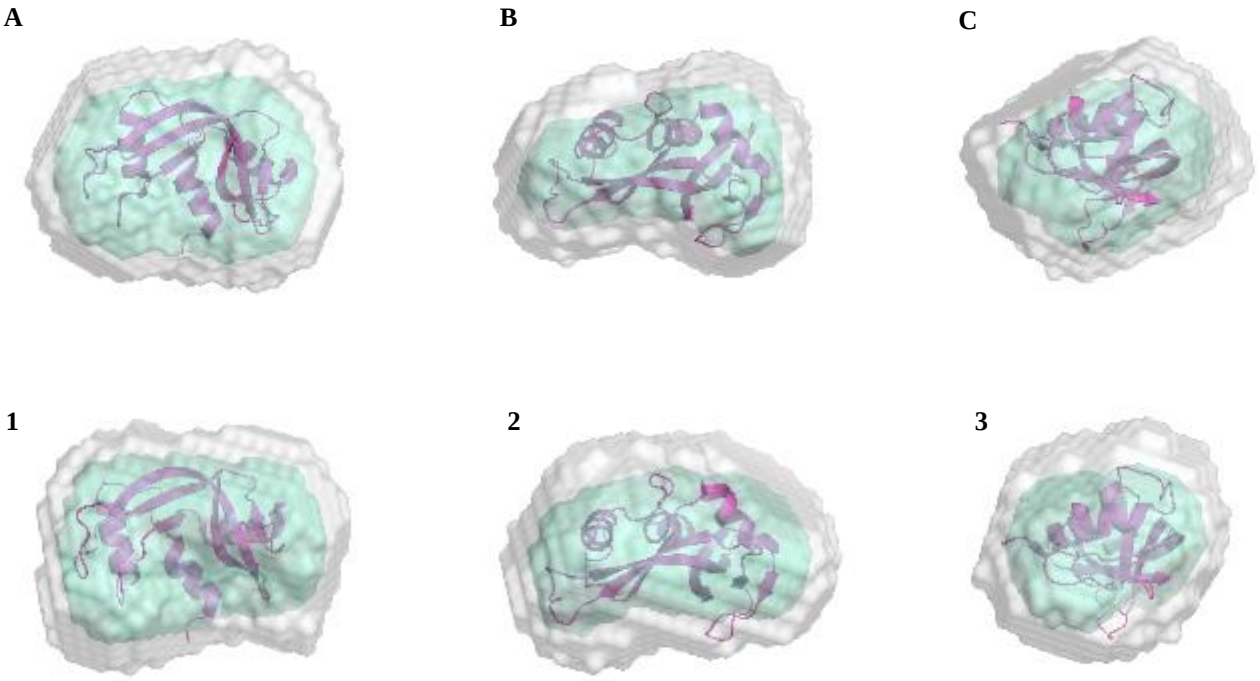


Figure 3.43. Refined and filtered molecular envelopes of ribonuclease A with and without LDAO detergent. The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope, the grey as the computed probability map generated by DAMAVER (Volkov & Svergun, 2003) and the pink ribbon cartoon is the known crystal structure of ribonuclease A (1FS3) taken from the protein data bank A, B, C Final molecular envelopes for ribonuclease A at 7 mg/ml in 0.3M NaCl, 50mM tris base, pH8 with front, side and end on views. 1, 2, 3Final molecular envelopes for ribonuclease A at 7 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8 shown with front, side and end on views.

3.10.5.7. Effects of LDAO upon the ribonuclease A molecular envelope

In order to deduce the effects of LDAO scattering upon the ribonuclease A models a final envelope for lysozyme in LDAO and without detergent was created. These DAMFILT (Volkov & Svergun, 2003) envelopes are shown in figure 3.44. The envelopes were overlaid alongside the surface structure from published X-ray diffraction experiments (PDB entry 1RNC – Aguilar et al., 1992). The envelopes correlate well and no major changes to the model can be identified upon the addition of LDAO to the sample and the blank.

The data analysis of the second control protein ribonuclease A was also promising, the radius of gyration was not seen to change between the two data both with and without the detergent LDAO. The experimentally determined radius of gyration for ribonuclease A both with and without LDAO was 15.7Å, this is larger than literature data of 14.3Å (Wang et al., 2008) however the radius of gyration can be effected by experimental conditions including temperature and pH (table 3.11).

Most importantly it was determined that the scattering caused by the detergent LDAO could be compensated for by the acquisition of good quality data and appropriate background subtraction. The control experiments also identified the need for extra care when preparing blanks with a dialysis matched blank being the most desired option.

Molecule	Radius of gyration (Å)
Ribonuclease A in 0.3M NaCl, 50mM tris base, pH8	15.71
Ribonuclease A in 0.1% v/v LDAO, 0.3M NaCl, 50mM tris base, pH8	15.68
Ribonuclease A literature data (Wang et al. 2008)	14.3

Table 3.11. A summary of the radius of gyration of ribonuclease A. Calculated *Rg* of ribonuclease A both with and without 0.1% v/v LDAO alongside the literature data. The calculated *Rg* with and without detergent are extremely close and both are relative to published literature values.

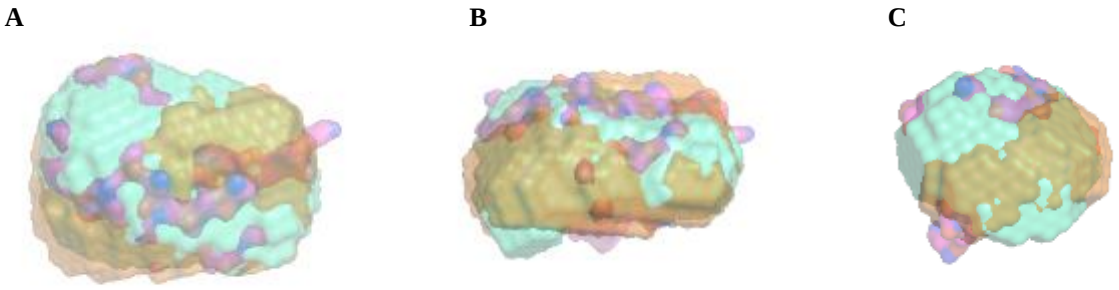


Figure 3.44. Assessment of the final molecular envelope of ribonuclease A by overlaying the models. Overlaid models of ribonuclease A generated in Pymol (Schrödinger, 2010) and showing the molecular envelope of ribonuclease A with LDAO (orange) and without (green), plus a surface representation of the ribonuclease A crystal structure (purple) taken from the protein data bank (1FS3). 1-2. Front and rear view. 3-4. End on views. 5-6. Side view of the overlaid envelopes.

3.10.6. When detergent scattering is not accounted for

In order to determine the importance of correct background subtraction, experiments were conducted where non-corresponding background scattering was subtracted from the sample scattering. Buffer blank scattering data from a sample of 0.3M NaCl, 50mM tris base at pH8 was subtracted from the scattering generated by a sample of lysozyme at 7 mg/ml and a sample of ribonuclease A at 5 mg/ml reconstituted in the above buffer plus 0.1% v/v LDAO. The resulting data shows that when detergent isn't accounted and isn't subtracted from the sample, ab inito modelling leads to the creation of micelle structures as shown in figure 3.46.

For these test experiments all analysis, modelling and refinement was carried out as previously described, the only difference was that LDAO was not accounted for within the background subtraction. The full parameters are shown in table 3.12 and the Guinier approximation, Kratky analysis and P(r) functions shown in figure 3.45.

Calculated Parameters	Lysozyme	Ribonuclease A
Rmax	44	44
Auto*Rg*	15.61 +-0.42 s*Rg* limits 0.511-1.288	15.67+-0.67 s*Rg* limits 0.6-1.32
Guinier *Rg*	16.2 +-0.161 s*Rg* limits 0.506-1.49	15.6 +-0.271 s*Rg* limits 0.488-1.22
Reciprocal space *Rg*	16.01	16.04
Real space *Rg*	15.99	16.03
I (0)	0.4036E+02	

Table 3.12. Parameters used for the analysis of the effects caused by poor background subtraction on the control protein lysozyme. All analysis was performed as previously described, a gnom output file (Svergun, 1992) was generated and used for molecular envelop generation.

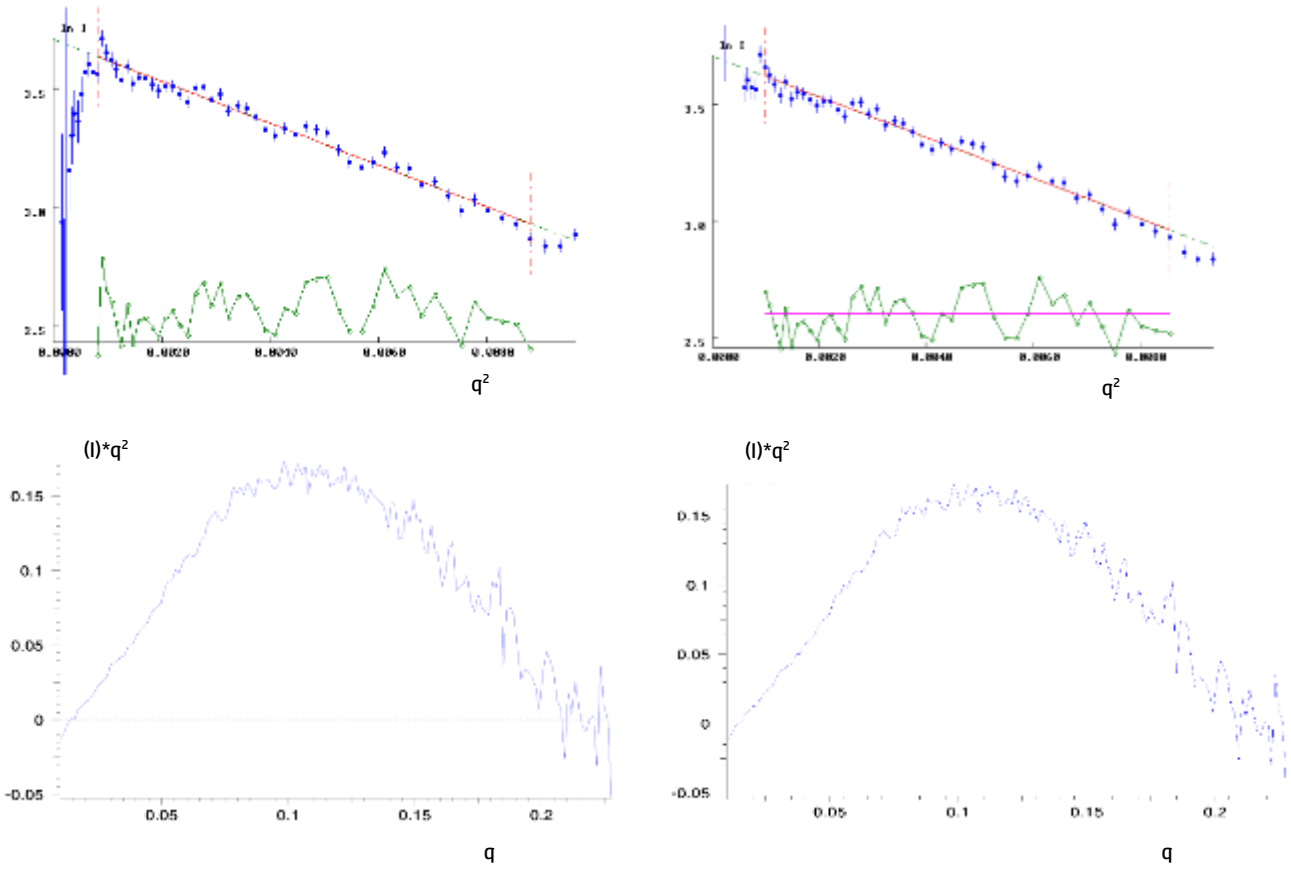


Figure 3.45. Graphical output from the analysis of lysozyme at 7 mg/ml and ribonuclease at 7 mg/ml with incorrect background subtraction. A) Guinier plot showing no obvious signs of problems, the Guinier approximation gave an estimated *Rg* of 16.2. B) A Kratky analysis with a typical bell shape curve indicating a folded protein. C) P(r) analysis of the sample suggesting a globular compact molecule. From initial analysis the poor back subtraction cannot be identified.

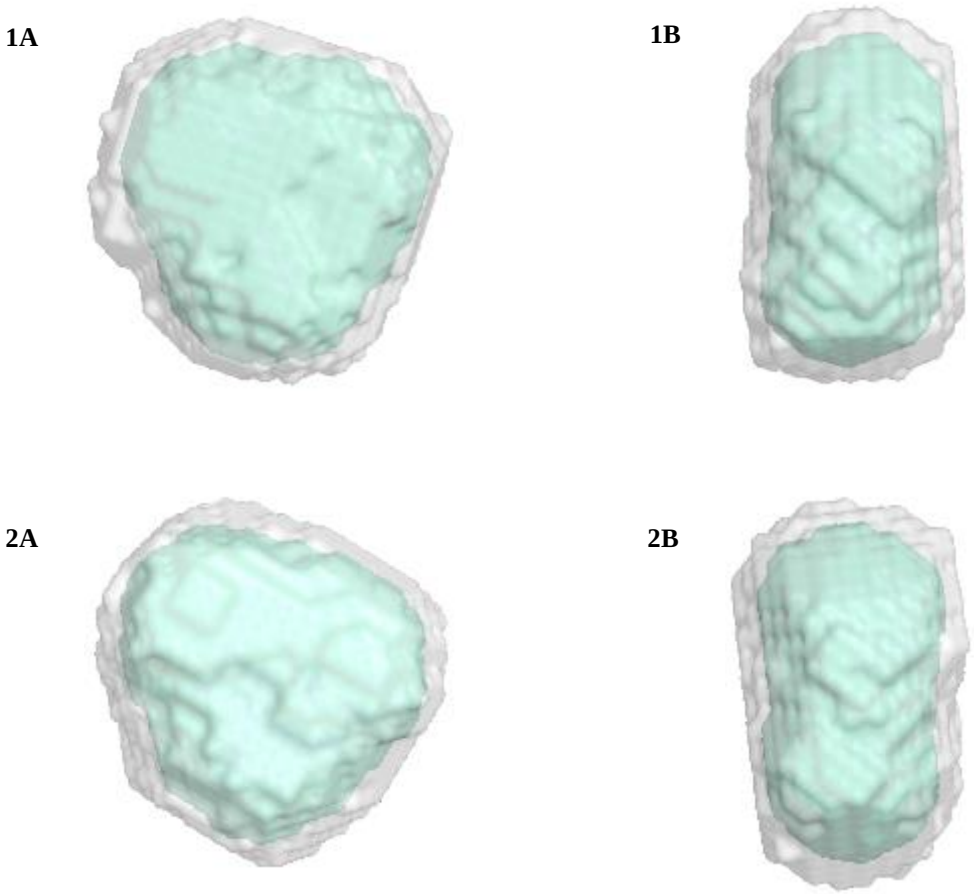


Figure 3.46. Final molecular envelopes for lysozyme and ribonuclease samples at 7 mg/ml with inappropriate background subtraction. Following the generation of 20 DAMMIF (Franke & Svergun, 2009) models for both lysozyme and ribonuclease A the final envelopes were determined after filtering and refinement as previously described. The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope and the grey the computed probability map generated by DAMAVER (Volkov & Svergun, 2003). The final models for both samples are representative of a LDAO micelle. By applying an incorrect background subtraction the contrast between the protein sample and the solvent is poor allowing visualisation of the target protein.

3.10.7.1. SAXS data collection of BAPKO_0422

SAXS data for BAPKO_0422 was collected for three different concentrations with the poly-histidine tag and for two data sets at different concentrations with the tag removed with the background subtracted scattering obtained shown in figure 3.47. Data was collected for 2400 seconds per frame, with ten frames for every sample and the blank. Various concentrations were used to ensure there was no relationship between an increasing the concentration and changes in the radius of gyration. Monodisperse systems from good quality samples devoid of protein aggregation or interpacticle interference should give a consistent radius of gyration regardless of the protein concentration. An increasing radius of gyration when increasing concentration is a sign of protein aggregation and is often seen by an upturn during Guinier analysis, in contrast a decreasing radius of gyration with an increase in sample concentration highlights interpacticle interference (Jacques and Trewhella 2009). A small upturn was noticed for all samples however this was also observed for the blank sample and decided that these first 12 points were due to the beamstop positioning for the NanoStar. Visual inspection of all 10 frames showed the level of scattering was consistent throughout the experiment and the protein appeared to remain stable for the full exposure with no signs of radiation damage.

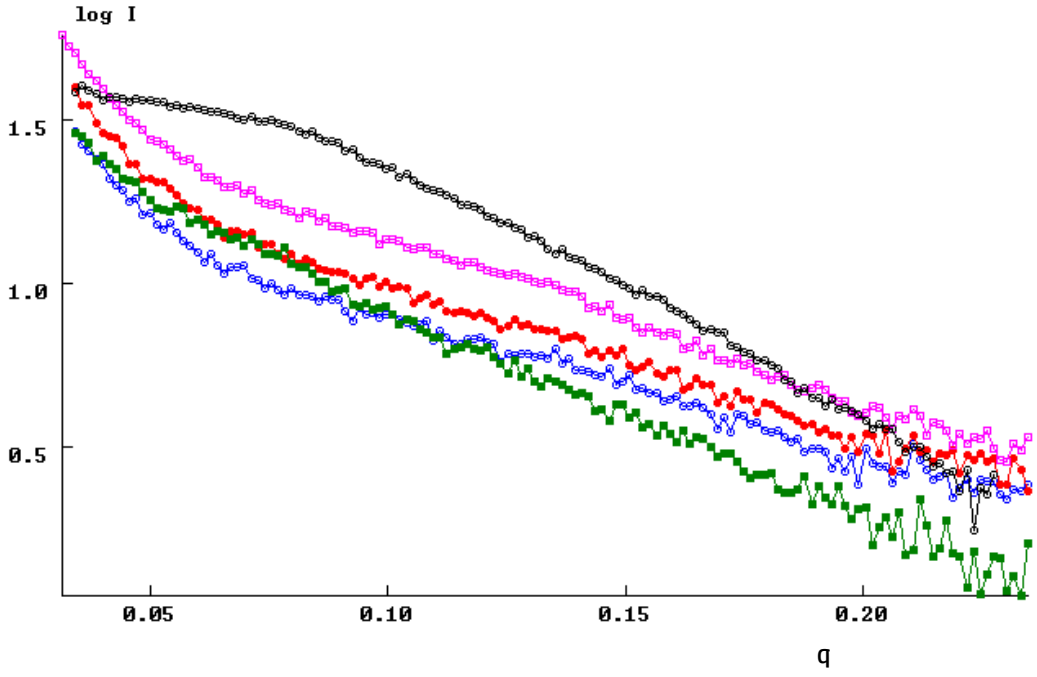


Figure 3.47. Background subtracted scattering data of samples of BAPKO_0422 at various concentrations. BLUE – Scattering data for tagged BAPKO_0422 at 2.5 mg/ml. RED – Scattering data for tagged BAPKO_0422 at 4 mg/ml.

PINK – Scattering data for tagged BAPKO_0422 at 4.5 mg/ml. GREEN – Scattering data for untagged BAPKO_0422 at 4 mg/ml. BLACK – Scattering data for untagged BAPKO_0422 at 3 mg/ml.

3.10.7.2. Calculating the radius of gyration of BAPKO_0422

All data analysis was carried out as previously described for the control protein lysozyme. The *Rg* for BAPKO_0422 was calculated using the Guinier approximation (figure 3.48) and also calculated in real space using GNOM (Svergun, 1992). Auto*Rg* was used where possible and where s*Rg* limits permitted. All *Rg* data for BAPKO_0422 is compiled in table 3.13. From initial analysis of the six data sets only three were suitable for modelling and the selection was based upon the quality of the data. Three data sets were excluded, two with the tag present and one without, these were discarded due to varying levels of aggregation identified from an upturn at low q^2 and due to an inconsistent Kratky plot indicating a mixture of folded/unfolded protein which is shown in figure 3.49.

3.10.7.3. Kratky analysis of BAPKO_0422

A Kratky plot or $I(q)*q^2$ versus s was produced for all BAPKO_0422 data sets using sasplot within the Primus package (Konarev et al., 2003). Kratky plots for all data sets are shown in figure 3.49. Two of the data sets for tagged BAPKO_0422 at 2.5 and 4 mg/ml produced a curve with a clear maximum but also with an increasing tail, the plots were representative of a partially folded protein indicating poor sample quality and the data sets were discarded. The three remaining data sets one for the histidine tagged protein and two for generated from untagged protein samples gave rise to the typical bell-shaped curve with a clear maximum value and these data sets were used for GNOM (Svergun, 1992) analysis and further modelling. Included and excluded data sets with their corresponding radius of gyration values are shown in table 3.12.

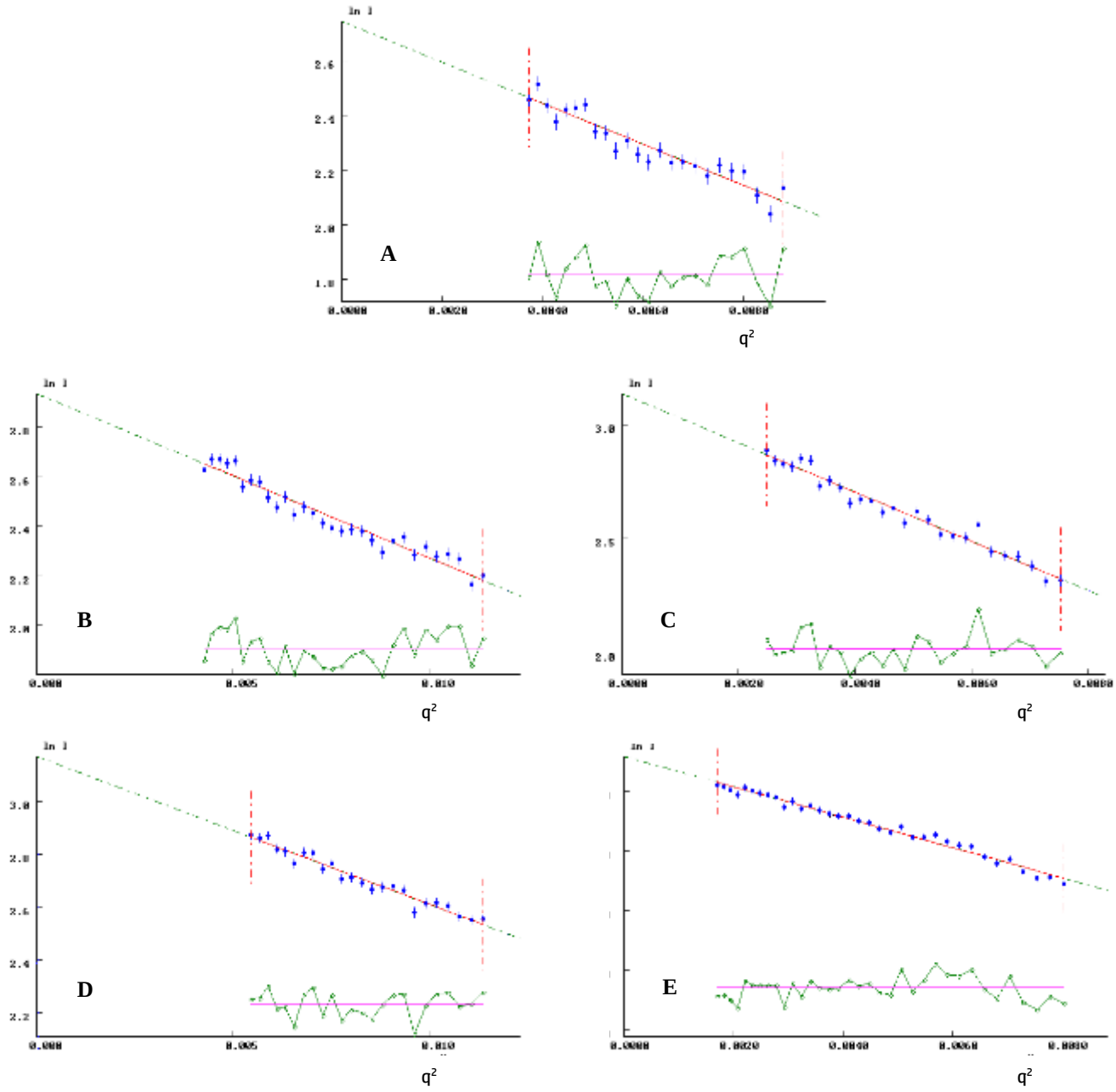


Figure 3.48. Guinier plots of BAPKO_0422. Guinier plots generated using Primus (Konarev et al., 2003) and representing data sets used for modelling. The scattering data is plotted as In[I(q)] vs q^2. The blue data points represent the scattering data, the red line shows the Guinier fit and the green data shows the corresponding residuals. A) Guinier analysis of tagged BAPKO_0422 at 2.5 mg/ml. B) Guinier plot of tagged BAPKO_0422 at 4 mg/ml. C) Guinier plot of tagged BAPKO_0422 at 4.5 mg/ml. D) Guinier plot of untagged BAPKO_0422 at 4 mg/ml. E) Guinier plot of untagged BAPKO_0422 at 3 mg/ml.

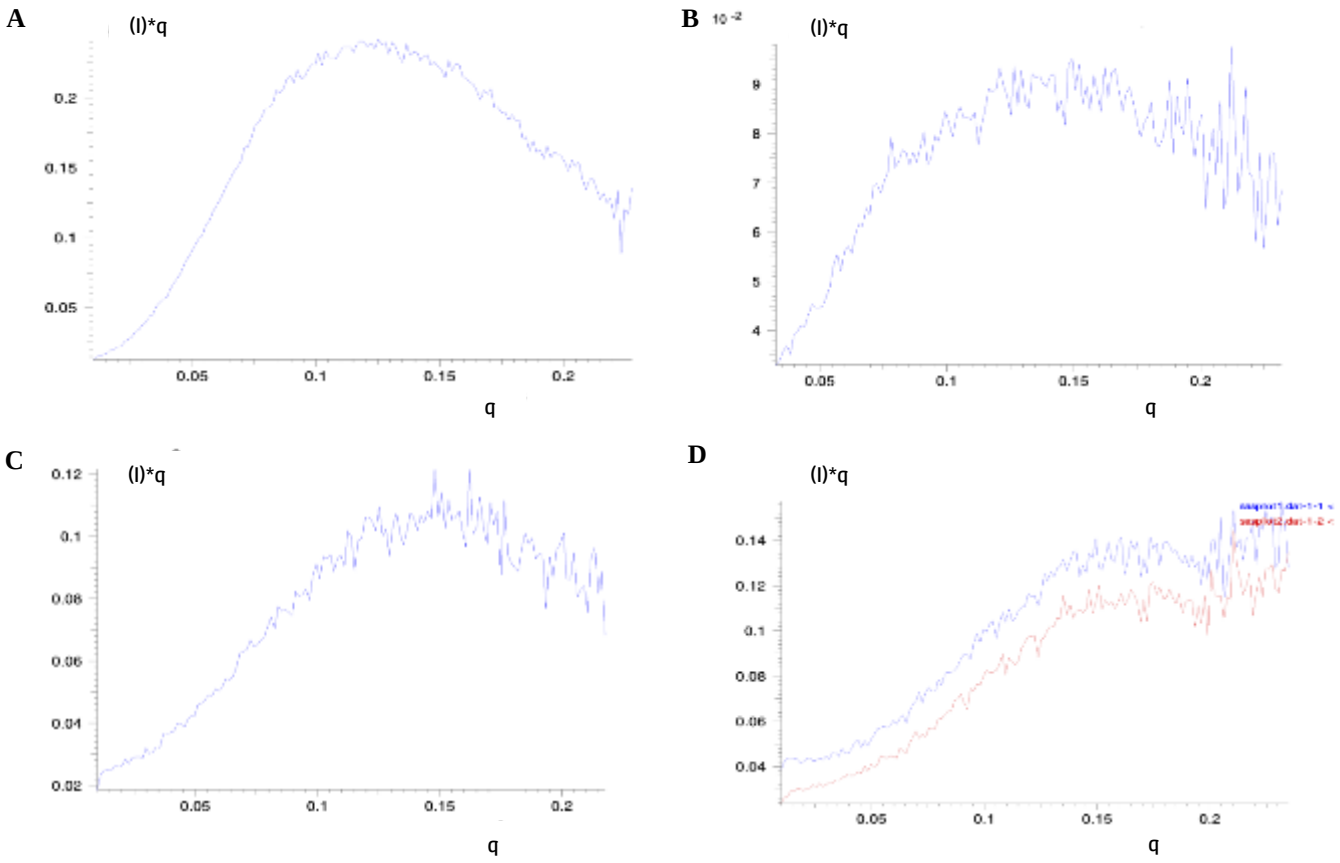


Figure 3.49. Kratky plot analysis of BAPKO_0422. Kratky plots or I(q)*q^2 plots of BAPKO_0422 samples generated using Primus (Konarev et al., 2003) A) Kratky plot of the scattering data generated by tagged BAPKO_0422 at 4.5 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. B) Kratky plot generated from 3 mg/ml untagged BAPKO_0422 in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. C) Kratky analysis of untagged BAPKO_0422 at 4 mg/ml. Plots A, B and C are representative of a folded globular like protein and these data sets were used for further analysis. D) Kratky analysis of the two excluded data sets, the Kratky represents partially folded protein and therefore the data was excluded. BLUE – BAPKO_0422 untagged at 4 mg/ml. RED – BAPKO_0422 histidine tagged at 2.5 mg/ml.

Sample	Auto *Rg* (Å)	Guinier *Rg* (Å)	Reciprocal space *Rg* (Å)	Real space *Rg* (Å)	Data used/ Discarded?
BAPKO_0422 tagged 2.5 mg/ml	----	15+-0.361 s*Rg* 0.916-1.41	16.78	16.79	Discarded
BAPKO_0422 tagged 4 mg.ml	12.48+-0.39 S*Rg* 0.959-1.278	14.1+-0.204 s*Rg* 0.922-1.5	16.75	16.77	Discarded
BAPKO_0422 tagged 4.5 mg/ml	13.91+-0.39 s*Rg* 0.789-1.65	13+-0.227 s*Rg* 0.958-1.38	16.62	16.65	Used
BAPKO_0422 untagged 4 mg/ml	----	12.2+-0.39 s*Rg* 0.896-1.50	14.4	14.47	Used
BAPKO_0422 untagged 3 mg/ml	11.68+-0.16 s*Rg* 0.482-0.931	12.4+-0.114 s*Rg* 0.510-1.11	14.29	14.28	Used

Table 3.13. Radius of gyration values for all BAPKO_422 data sets. The radius of gyration was calculated for all data sets using the Primus (Konarev et al., 2003) Auto*Rg* function and by manual Guinier approximation. The values were further calculated using the full experimental data both in reciprocal and real space during GNOM (Svergun, 1992) analysis. Following Kratky analysis data sets, 2.5 and 4 mg/ml tagged BAPKO_0422 were discarded due to the presence of partially folded protein.

3.10.7.4. GNOM analysis and evaluation of the pair-distance distribution function P(r) of BAPKO_0422

Following the initial analysis of the background subtracted scattering data GNOM (Svergun, 1992) was used as an indirect transform program for examination of the P(r) function and for preparation of a gnom.out file for use in ab initio modelling. The three BAPKO_0422 data sets were submitted to GNOM (Svergun, 1992) independently where their radius of gyration was calculated in both reciprocal and real space, shown in figure 3.50. The P(r) function for all three data sets was representative of a single domain compact protein.

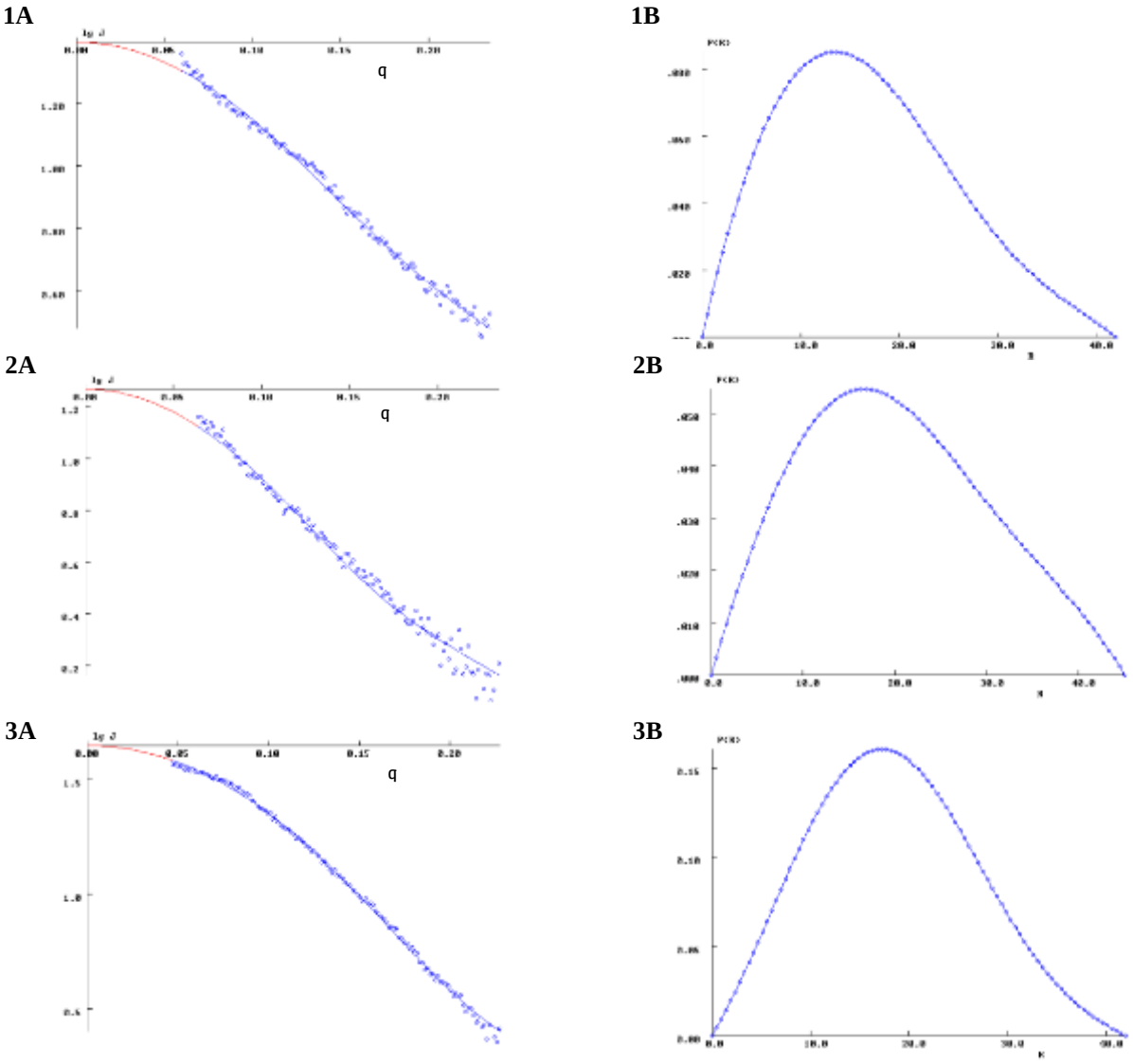


Figure 3.50. GNOM analysis of the BAPKO_0422 data sets. GNOM (Svergun, 1992) was used as an indirect transform tool, the radius of gyration was calculated for all data sets in both real and reciprocal space. **1-3A** – Reciprocal space fit for tagged BAPKO_0422 at 3 mg/ml, untagged BAPKO_0422 at 4 mg/ml and tagged BAPKO_0422 at 4.5 mg/ml respectively. **1-3B** – P(r) function for tagged BAPKO_0422 at 3 mg/ml, untagged BAPKO_0422 at 4 mg/ml and tagged BAPKO_0422 at 4.5 mg/ml respectively.

3.10.7.5. Generating molecular envelopes for untagged BAPKO_0422

Using the same methodology as was used for the control protein lysozyme, data for untagged BAPKO_0422 was transformed using GNOM (Svergun, 1992) and the output file submitted to DAMMIF (Franke & Svergun, 2009) for ab initio modelling. The process was repeated until twenty impendent molecular envelopes were produced for each data set giving a total of forty models. The initial envelopes for both data sets are shown in figure 3.51.

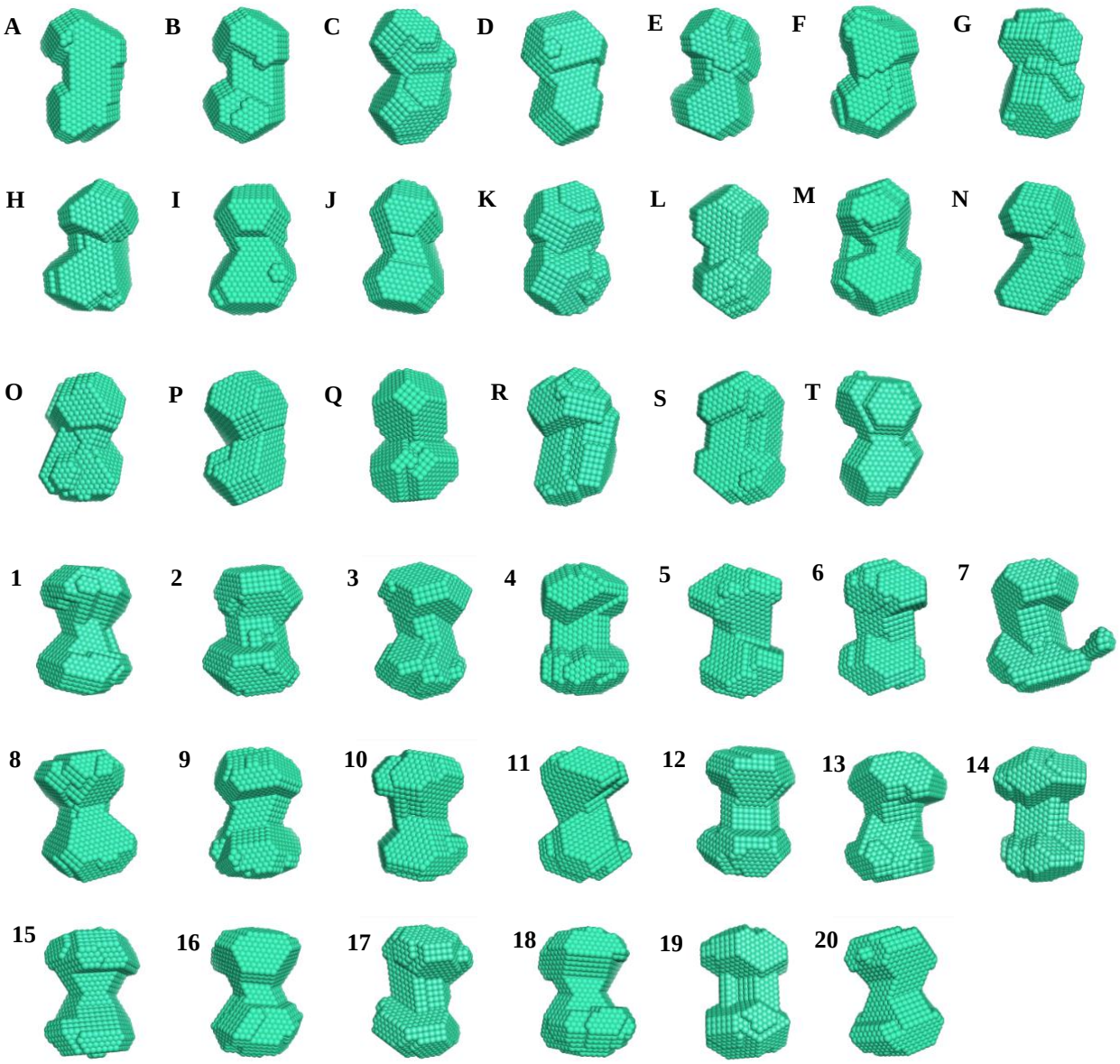


Figure 3.51. DAMMIF molecular envelopes of untagged BAPKO_0422. Molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models per data set.

Envelopes A-T were generated from scattering data from untagged BAPKO_0422 at 4.5 mg/ml. Envelopes 1-20 generated from scattering data from untagged BAPKO_0422 at 3 mg/ml.

3.10.7.6. Evaluation and refinement of untagged BAPKO_0422

For initial modelling both data sets of untagged BAPKO_422 at 4.5 and 3 mg/ml were refined independently from each other. Assessment of the generated envelopes was carried out by DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003) as previously described for the control proteins. Outliers were identified by two points of standard deviation from the calculated NDS mean and were removed. Envelopes P and 20 were chosen by DAMSEL (Volkov & Svergun, 2003) as the reference envelopes for superimposition with full DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003) data summarised in table 3.14. Assessment by eye of the DAMMIF (Franke & Svergun, 2009) envelopes highlighted that model 7 generated for untagged BAPKO_0422 at 3 mg/ml was rather different from the other models, DAMSEL (Volkov & Svergun, 2003) confirmed this and removed the envelope alongside model S for the 4 mg/ml data.

Following refinement DAMAVER (Volkov & Svergun, 2003) was used to produce an averaged probability map this map was then filtered by DAMFILT (Volkov & Svergun, 2003) which removed regions of low occupancy to give a final molecular envelope for both of the untagged BAPKO_0422 data sets which are shown in figure 3.52.

E. coli, OmpX was selected as an example OMP to be docked into the experimentally derived molecular envelope. This protein was used to make initial assessments of the SAXS model and was selected as it represented a well-studied 8-stranded beta barrel. OmpX was selected over OmpA as its dimensions were proposed to be most similar to the Borrelia OMPs. The relative order of OmpX was hypothesised to represent the Borrelia OMPs. OmpX is composed of an 8-stranded beta barrel which has an extensive hydrogen bonding network of conserved residues and resembles an inverse micelle. The extracellular loops are much more variable than the barrel region and much larger than the intracellular turns. The docked structure is represented as a cartoon of the secondary structure elements within the molecular envelope as shown in figure 3.52.

Solution	DAMSEL NSD average (Standard deviation of NSD = 0.010 and 0.048)	Include/Discard?	DAMSUP NSD superimposition value	Indicative fit (Chi squared)
Untagged BAPKO_0422 at 4.5 mg/ml				
A	0.517	**Include**	0.492	4.363
B	0.515	**Include**	0.497	4.362
C	0.528	**Include**	0.522	4.363
D	0.500	**Include**	0.506	4.363
E	0.503	**Include**	0.510	4.363
F	0.502	**Include**	0.446	4.362
G	0.513	**Include**	0.516	4.362
H	0.505	**Include**	0.482	4.363
I	0.511	**Include**	0.524	4.363
J	0.508	**Include**	0.429	4.364
K	0.515	**Include**	0.494	4.363
L	0.502	**Include**	0.486	4.363
M	0.521	**Include**	0.501	4.362
N	0.530	**Include**	0.506	4.363
O	0.515	**Include**	0.537	4.363
P	0.497	**Reference**	0.000	4.363
Q	0.513	**Include**	0.457	4.364
R	0.513	**Include**	0.527	4.363
S	0.535	**Discard**	-----	4.363
T	0.518	**Include**	0.520	4.363
Untagged BAPKO_0422 at 3 mg/ml				
1	0.480	**Include**	0.487	3.953
2	0.491	**Include**	0.479	3.960
3	0.489	**Include**	0.472	3.962
4	0.482	**Include**	0.489	3.957
5	0.492	**Include**	0.489	3.954
6	0.492	**Include**	0.494	3.953
7	0.702	**Discard**	-----	3.962
8	0.496	**Include**	0.496	3.962
9	0.488	**Include**	0.498	3.958
10	0.478	**Include**	0.256	3.955
11	0.488	**Include**	0.492	3.957
12	0.480	**Include**	0.442	3.957
13	0.492	**Include**	0.476	3.958
14	0.480	**Include**	0.438	3.958
15	0.497	**Include**	0.485	3.960
16	0.482	**Include**	0.428	3.958
17	0.483	**Include**	0.488	3.958
18	0.488	**Include**	0.471	3.960
19	0.499	**Include**	0.494	3.960
20	0.478	**Reference**	0.000	3.957

Table 3.14. DAMSEL and DAMSUP scoring for untagged BAPKO_0422. A summary of the evaluation and refinement of untagged BAPKO_0422 models by DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003). DAMSEL (Volkov & Svergun, 2003) determined the mean NSD as 0.513 and 0.498 for the 4.5 mg/ml and 3 mg/ml data sets respectively. Envelope S and 7 were calculated as outliers and the data discarded. Data for envelopes P and 20 were identified as the most probably solutions in each case and this data was used as the reference set for superimposition. DAMSUP (Volkov & Svergun, 2003) NSD values were calculated relative to the reference set. The chi squared value represents the indicative fit calculated by DAMMIF (Volkov & Svergun, 2003) for each generated model.

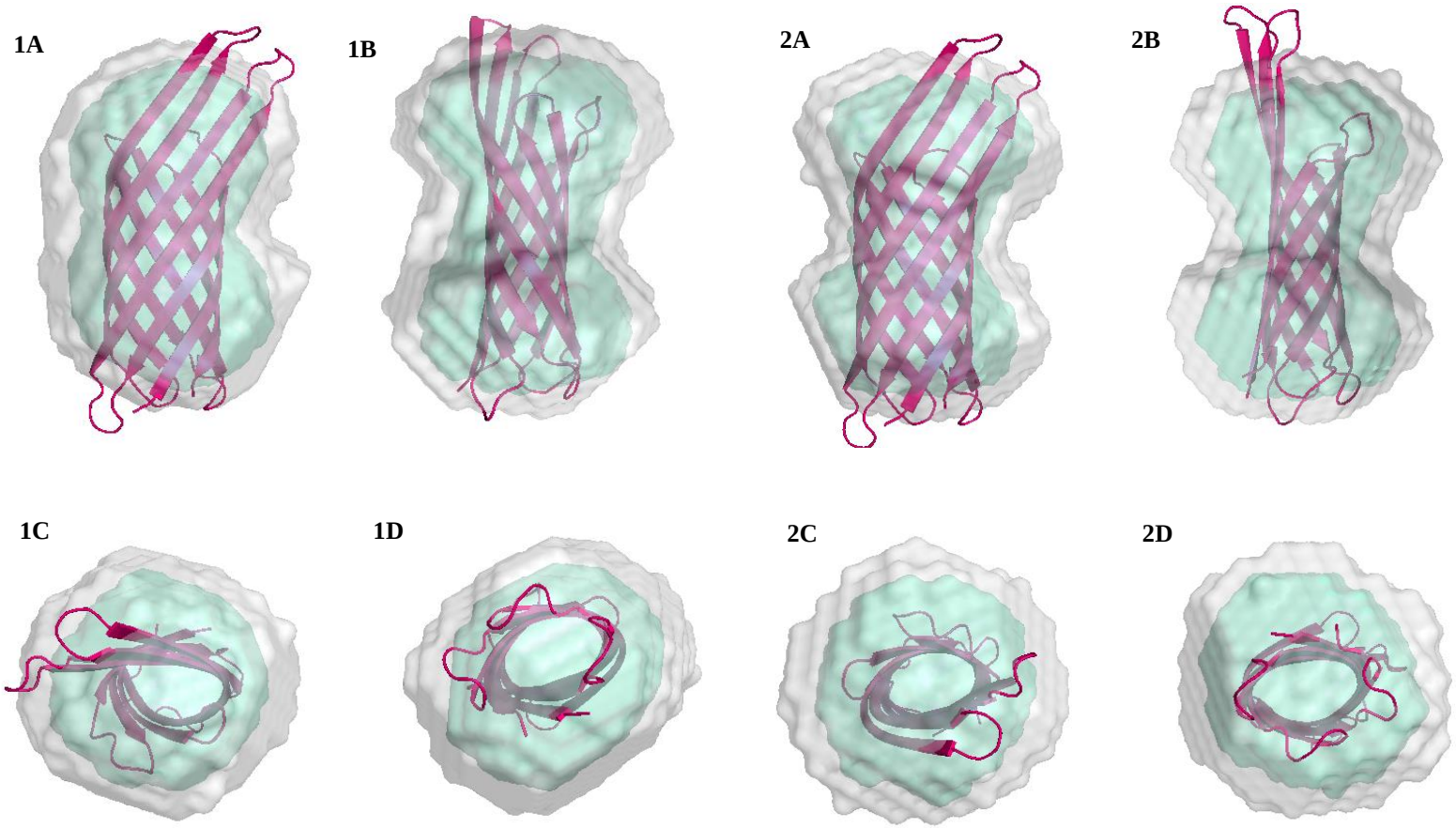


Figure 3.52. Final molecular envelopes for untagged BAPKO_0422. Following the generation of 20 DAMMIF (Franke & Svergun, 2009) models the final envelopes were determined after filtering and refinement as previously described. The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope, the grey as the computed probability map generated by DAMAVER (Volkov & Svergun, 2003) and the pink ribbon cartoon is the known crystal structure of *E. coli* OmpX (1QJ8 - Vogt & Schulz, 1999) taken from the protein data bank. All models were generated using Pymol (Schrödinger, 2010). 1A-D) Final molecular envelope of untagged BAPKO_0422 at 4.5 mg/ml. 2A-D) The final model of untagged BAPKO_0422 at 3 mg/ml.

3.10.7.7. Generating molecular envelopes for histidine tagged BAPKO_0422

Molecular envelopes were generated for the final BAPKO_0422 data set which contained a polyhistidine tag, the output file generated from GNOM (Svergun, 1992) for the data set was submitted to DAMMIF (Franke & Svergun, 2009) for ab initio modelling. DAMMIF (Franke & Svergun, 2009) modelling was repeated twenty times to generate 20 independent envelopes for further evaluation and refinement. The envelopes for tagged BAPKO_0422 are shown in figure 3.53.

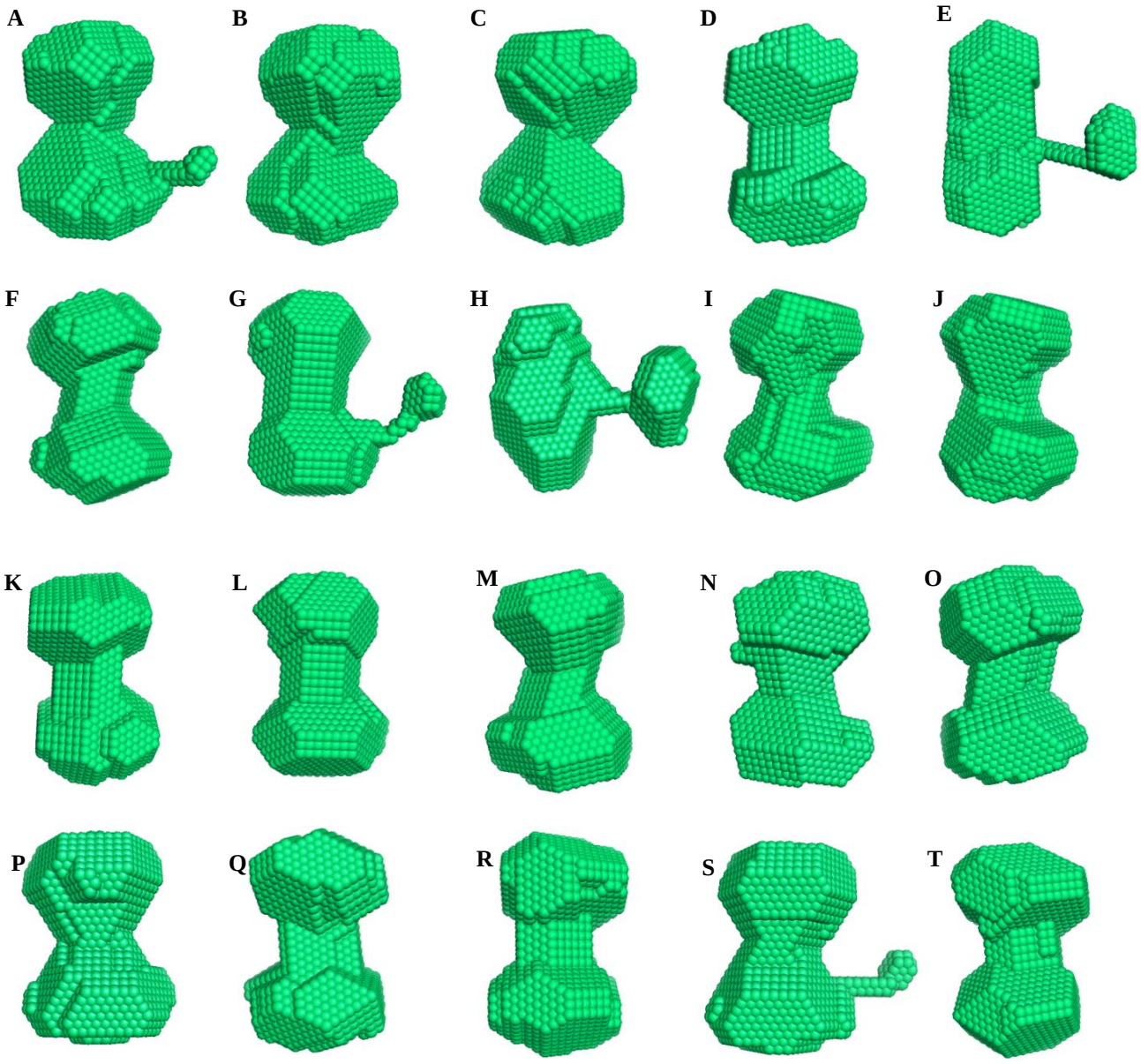


Figure 3.53. DAMMIF molecular envelopes of histidine tagged BAPKO_0422. Twenty envelopes were independently generated by DAMMIF (Franke & Svergun, 2009) for tagged BAPKO_0422 at 4 mg/ml which were then evaluated by DAMSEL (Volkov & Svergun, 2003).

3.10.7.8. Evaluation and refinement of histidine tagged BAPKO_0422

The envelopes generated by DAMMIF (Franke & Svergun, 2009) for histidine tagged BAPKO_0422 were assessed by DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003) as previously described. The NSD of envelope J was calculated as significantly different from the remaining envelopes and was removed before the files were submitted to DAMSUP (Volkov & Svergun, 2003). Envelopes D and O were observed to have the lowest NSD values and were automatically selected as the reference envelopes for superimposition as shown in table 3.15.

Following the removal of the outliers and the superimposition of all remaining models the coordinate and PDB files were submitted to DAMAVER (Volkov & Svergun, 2003) in order to compute a probability map which was later filtered by DAMFILT (Volkov & Svergun, 2003). The final DAMAVER (Volkov & Svergun, 2003) and DAMFILT (Volkov & Svergun, 2003) envelopes are shown in figure 3.54.

Solution	DAMSEL NSD average (Standard deviation of NSD = 0.128)	Include/Discard?	DAMSUP NSD superimposition value	Indicative fit (Chi squared)
Histidine tagged BAPKO_0422				
A	0.514	Include	0.465	4.750
B	0.519	Include	0.475	4.749
C	0.520	Include	0.496	4.751
D	0.491	Reference	0.000	4.750
E	0.528	Include	0.503	4.749
F	0.509	Include	0.454	4.751
G	0.528	Include	0.495	4.749
H	0.532	Include	0.508	4.750
I	0.524	Include	0.496	4.750
J	1.092	Discard	-----	4.750
K	0.518	Include	0.452	4.749
L	0.520	Include	0.453	4.750
M	0.520	Include	0.496	4.750
N	0.531	Include	0.503	4.750
O	0.491	Reference	0.000	4.749
P	0.526	Include	0.484	4.750
Q	0.516	Include	0.434	4.750
R	0.538	Include	0.530	4.749
S	0.530	Include	0.499	4.750
T	0.521	Include	0.497	4.750

Table 3.15. A summary of the evaluation and refinement of histidine tagged BAPKO_0422 models by DAMSEL and DAMSUP. DAMSEL (Volkov & Svergun, 2003) determined the mean NSD as 0.548. Envelope J was calculated as an outlier and the data discarded. Data for envelopes D and O were identified as the most probably solutions in each case and this data used as the reference set for superimposition. DAMSUP (Volkov & Svergun, 2003) NSD values were calculated relative to the reference set. The chi squared value represents the indicative fit calculated by DAMMIF (Volkov & Svergun, 2003) for each generated model.

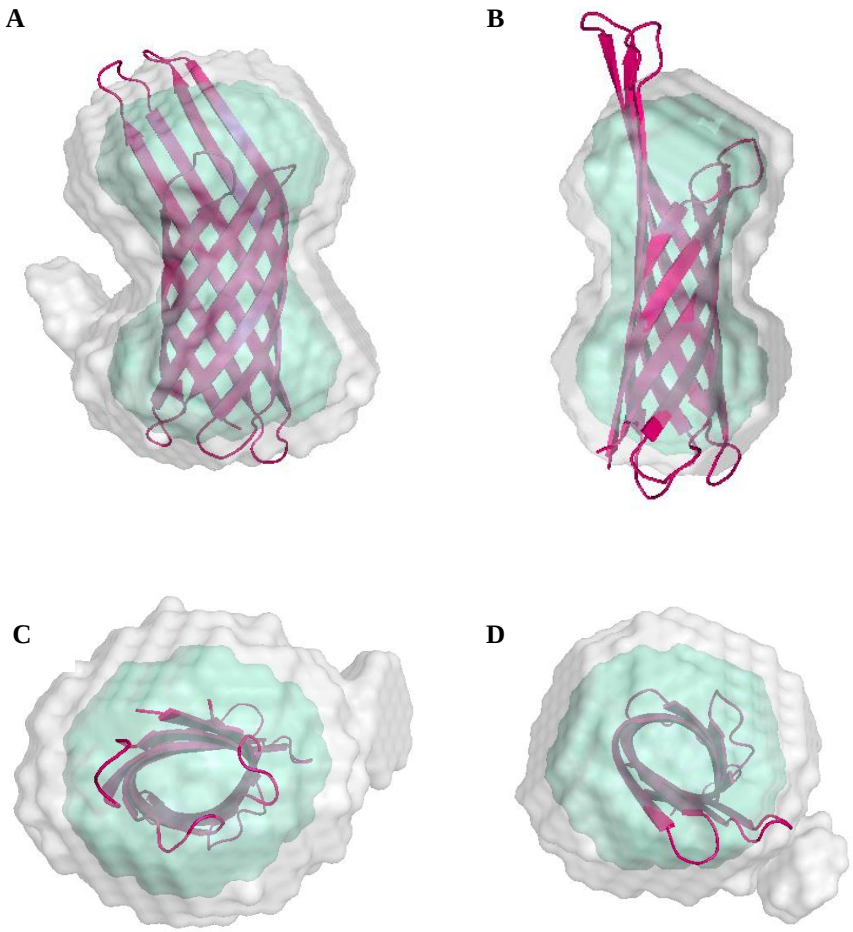


Figure 3.54. Refined and filtered molecular envelope for histidine tagged BAPKO_0422 at 4 mg/ml. Following the generation of 20 DAMMIF (Franke & Svergun, 2009) models the final envelope was determined after filtering and refinement by DAMAVER (Volkov & Svergun, 2003). The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope and the grey is the computed probability map generated by DAMAVER (Volkov & Svergun, 2003). The purple cartoon is OmpX from *E. coli*, data taken from the protein data bank (1QJ8 - Vogt & Schulz, 1999).

3.10.8.1. SAXS data collection for BB_0562

Scattering data for BB_0562 was collected over three different concentrations both with and without the poly-histidine tag. The scattering data following background subtraction is shown in figure 3.55. Scattering data was collected for 2400 seconds per frame with ten frames for every sample and for the blank.

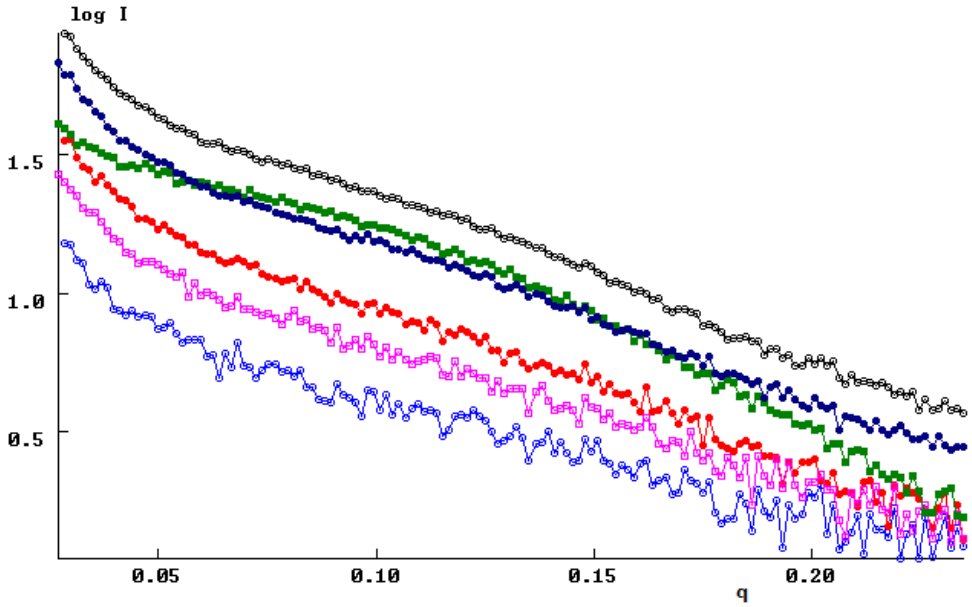


Figure 3.55. Background subtracted scattering data for samples of BB_0562 at various concentrations. From top curve to bottom. **A) BLACK** – Scattering data for tagged BB_0562 at 6.5mg/ml. **B) DARK BLUE** – Scattering data for tagged BB_0562 at 6mg/ml. **C) GREEN** – Scattering data for untagged BB_0562 at 5mg/ml. **D) RED** – Scattering data for tagged BB_0562 at 4mg/ml. **E) PURPLE** – Scattering data for untagged BB_0562 at 3mg/ml. **F) LIGHT BLUE** – Scattering data for untagged BB_0562 at 2.5mg/ml.

3.10.8.2. Calculating the radius of gyration of BB_0562

All data analysis was carried out as previously described for the control protein lysozyme. The radius of gyration for BB_0562 was calculated using the Guinier approximation and also calculated in real and reciprocal space using GNOM (Svergun, 1992). Auto*Rg* was used where possible and where s*Rg* limits permitted. All radius of gyration values for BB_0562 are compiled in table 3.16 and the Guinier analysis shown in figure 3.56.

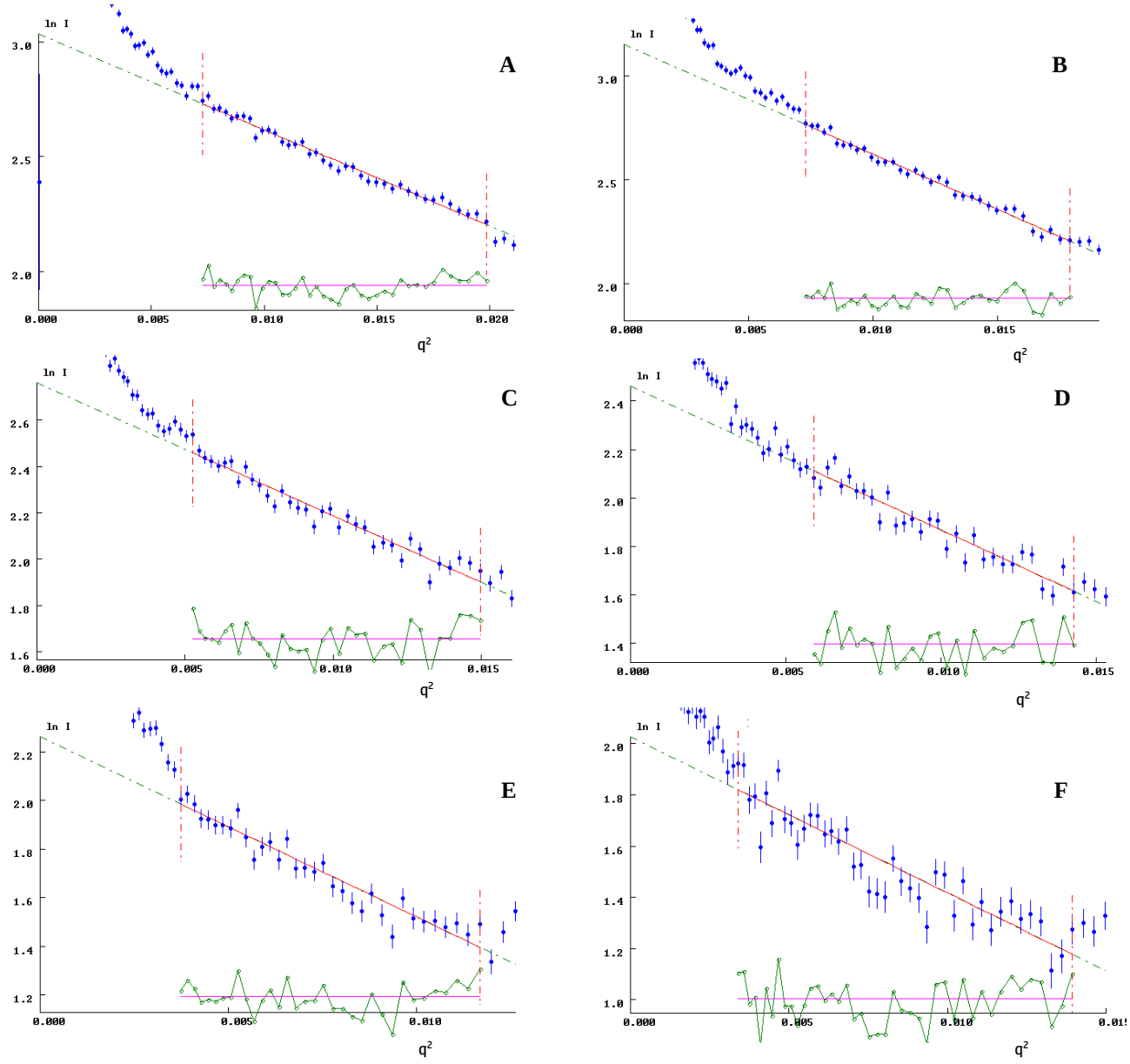


Figure 3.56. Manual Guinier plots for BB_0562. Guinier plots generated using Primus (Konarev et al., 2003) and representing all data sets. The scattering data is plotted as In[I(q)] vs q^2. The blue data points represent the scattering data, the red line shows the Guinier fit and the green data shows the corresponding residuals. A) Guinier analysis of tagged BB_0562 at 6.5 mg/ml. B) Guinier plot of tagged BB_0562 at 6 mg/ml. C) Guinier plot of untagged BB_0562 at 5 mg/ml. D) Guinier plot of tagged BB_0562 at 4 mg/ml. E) Guinier plot of untagged BB_0562 at 3 mg/ml. F) Guinier plot of untagged BB_0562 at 2.5 mg/ml. The Guinier region was particularly difficult to determine and the dataset discarded.

3.10.8.3. Kratky analysis of BB_0562

Kratky plots were generated for all six data sets for BB_0562 and are shown in figure 3.57. From both the Guinier approximation and the Kratky plots only three of the six data sets were deemed suitable for modelling. This selection was based upon the quality of the data with data sets A (tagged BB_0562 at 6.5 mg/ml), D (tagged BB_0562 at 4 mg/ml) and F (untagged BB_0562 at 2.5 mg/ml) being excluded due to the Kratky analysis which failed to show a bell shaped curve with a clear maximum peak which returned towards zero, this was indicative of a sample containing partially folded protein.

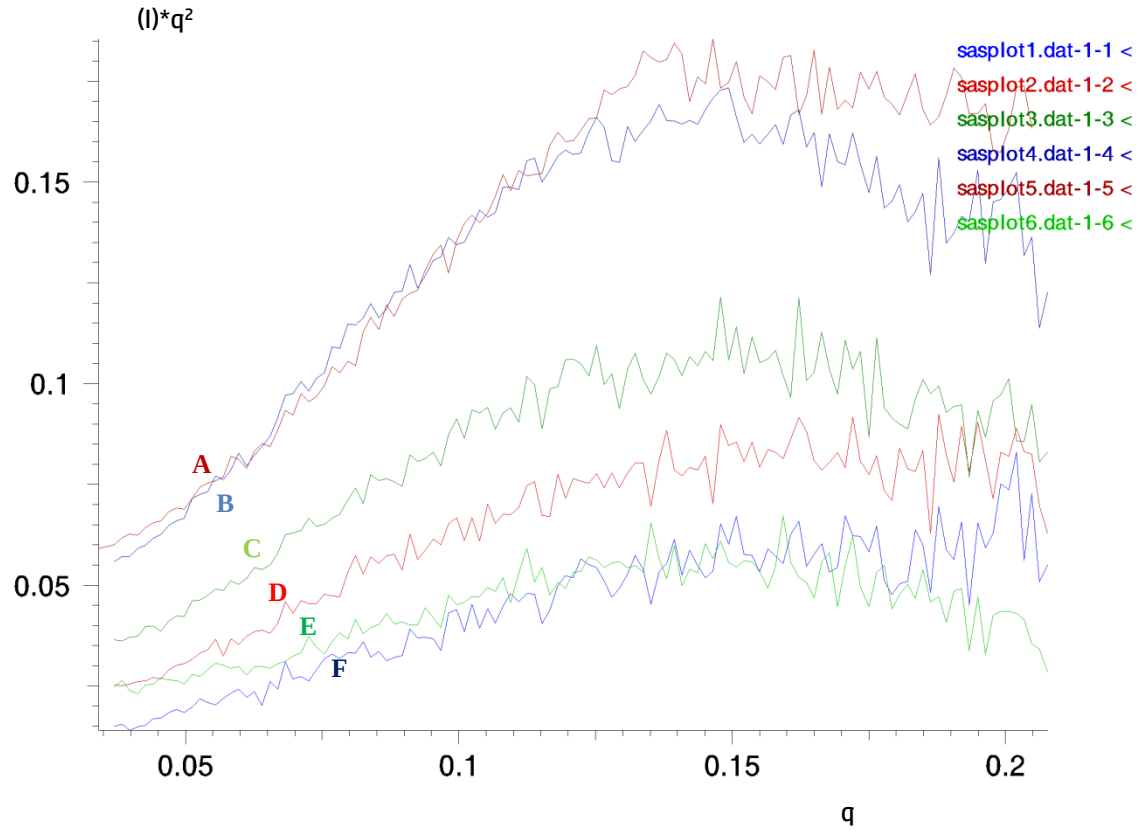


Figure 3.57. Kratky plot analysis of BB_0562. A) DARK RED Kratky plot or I(q)*q^2 plot of the scattering data generated by tagged BB_0562 at 6.5 mg/ml in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. B) BLUE - Kratky plot generated from a sample of tagged BB_0562 at 6 mg/ml. C) LIGHT GREEN - Kratky analysis of untagged BB_0562 at 5 mg/ml. D) RED - Kratky analysis of tagged BB_0562 at 4 mg/ml. E) GREEN – untagged BB_0562 at 3 mg/ml. F) DARK BLUE – untagged BB_0562 at 2.5 mg/ml. The Kratky analysis of samples of BB_0562 at 6 mg/ml, 5 mg/ml and 3 mg/ml, B/C/E respectively show a well behaved bell shape curve with a clear maxima. These samples were selected for further modelling.

From evaluation of both the Guinier fit and the Kratky analysis it was decided to exclude BB_0562 data generated for untagged BB_0562 at 2.5 mg/ml, and for tagged BB_0562 at 4 and 6.5 mg/ml. This data was discarded due to difficulty when applying the Guinier approximation and due to the generated Kratky plots shown in figure 3.57. Kratky analysis of the excluded data sets failed to show a bell shaped curve with a clear maximum peak which returned towards zero, this was indicative of a sample containing partially folded protein.

Sample	**Auto *Rg***	**Guinier *Rg***	**Rmax**	**Reciprocal space *Rg***	**Real space *Rg***	**Data used/ Discarded?**
A. BB_0562 6.5 mg/ml Tagged	-----	11.02+-0.109 s*Rg* 0.970-1.57	45	14.03	14.09	Discarded
B. BB_0562 6 mg/ml Tagged	----	14.3+-0.125 s*Rg* 0.915-1.61	45	14.27	14.31	Used
C. BB_0562 5 mg/ml Untagged	13.91+-0.39 s*Rg* 0.789-1.65	14.1+-0.158 s*Rg* 0.882-1.66	45	14.29	14.34	Used
D. BB_0562 4 mg/ml Tagged	38.57+-11.33 s*Rg*1.152-1.701	13.8+-0.213 s*Rg* 0.861-1.59	45	14.43	14.50	Discarded
E. BB_0562 3 mg/ml Untagged	14.82+-11.15 s*Rg* 0.946-1.286	14.6+-0.246 s*Rg* 0.872-1.64	45	14.07	14.09	Used
F. BB_0562 2.5 mg/ml Untagged	----	13.5+-0.294 s*Rg* 0.788-1.6	45	14.5	14.6	Discarded

Table 3.16. Radius of gyration values for all BB_0562 data sets. The radius of gyration was calculated for all data sets using the Primus (Konarev et al., 2003) Auto*Rg* function and by manual Guinier approximation. The values were further calculated using the full experimental data both in reciprocal and real space during GNOM (Svergun, 1992) analysis.

3.10.8.4. GNOM analysis and evaluation of the pair-distance distribution function of BB_0562

Following the removal of unsuitable data the two remaining data sets for untagged BB_0562 and the single dataset for tagged BB_0562 were processed by GNOM (Svergun, 1992) and the P(r) function evaluated. All parameters were kept as standard and an Rmax of 45 selected. The chosen Rmax reflected the point in which the P(r) function returned to zero with a steady decline. The radius of gyration for each of the samples was calculated both in reciprocal and reals space with the curves shown in figure 3.58.

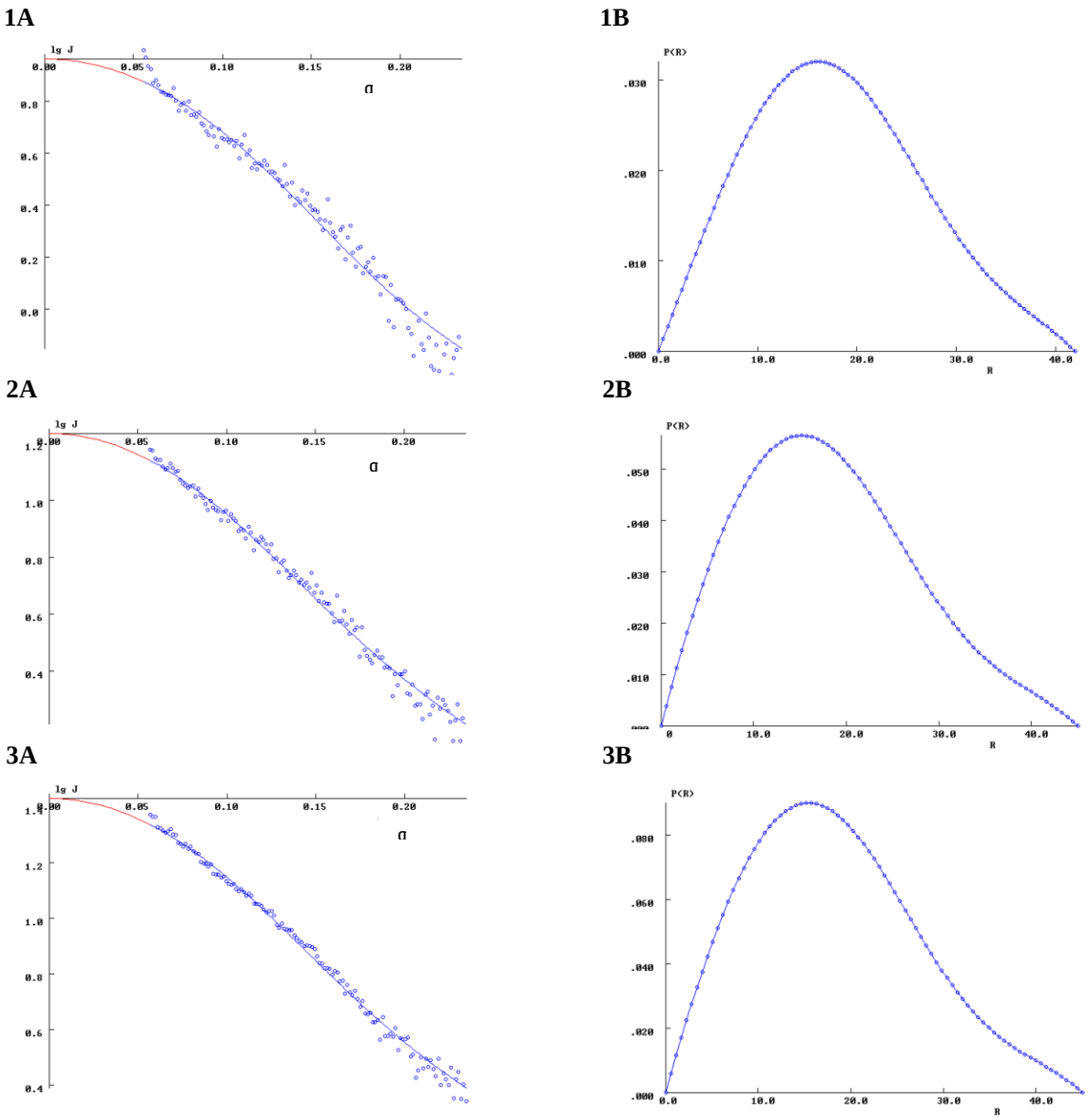


Figure 3.58. Calculation of reciprocal and real space *Rg* with P(r) analysis of included BB_0562 data sets. GNOM (Svergun, 1992) was used as an indirect transform tool, the radius of gyration was calculated for all data sets in both real and reciprocal space. **1-3A** – Reciprocal space fit for tagged BB_0562 at 6 mg/ml, untagged BB_0562 at 5 mg/ml and untagged BB_0562 at 3 mg/ml respectively. **1-3B** – P(r) function for tagged BB_0562 at 6 mg/ml, untagged BB_0562 at 5 mg/ml and untagged BB_0562 at 3 mg/ml respectively..

3.10.8.5. Generating molecular envelopes for untagged BB_0562

The two remaining data sets for untagged BB_0562 were used for ab initio modelling as previously described. Twenty independent models were generated by DAMMIF (Franke & Svergun, 2009) for each data set which are shown in figure 3.59.

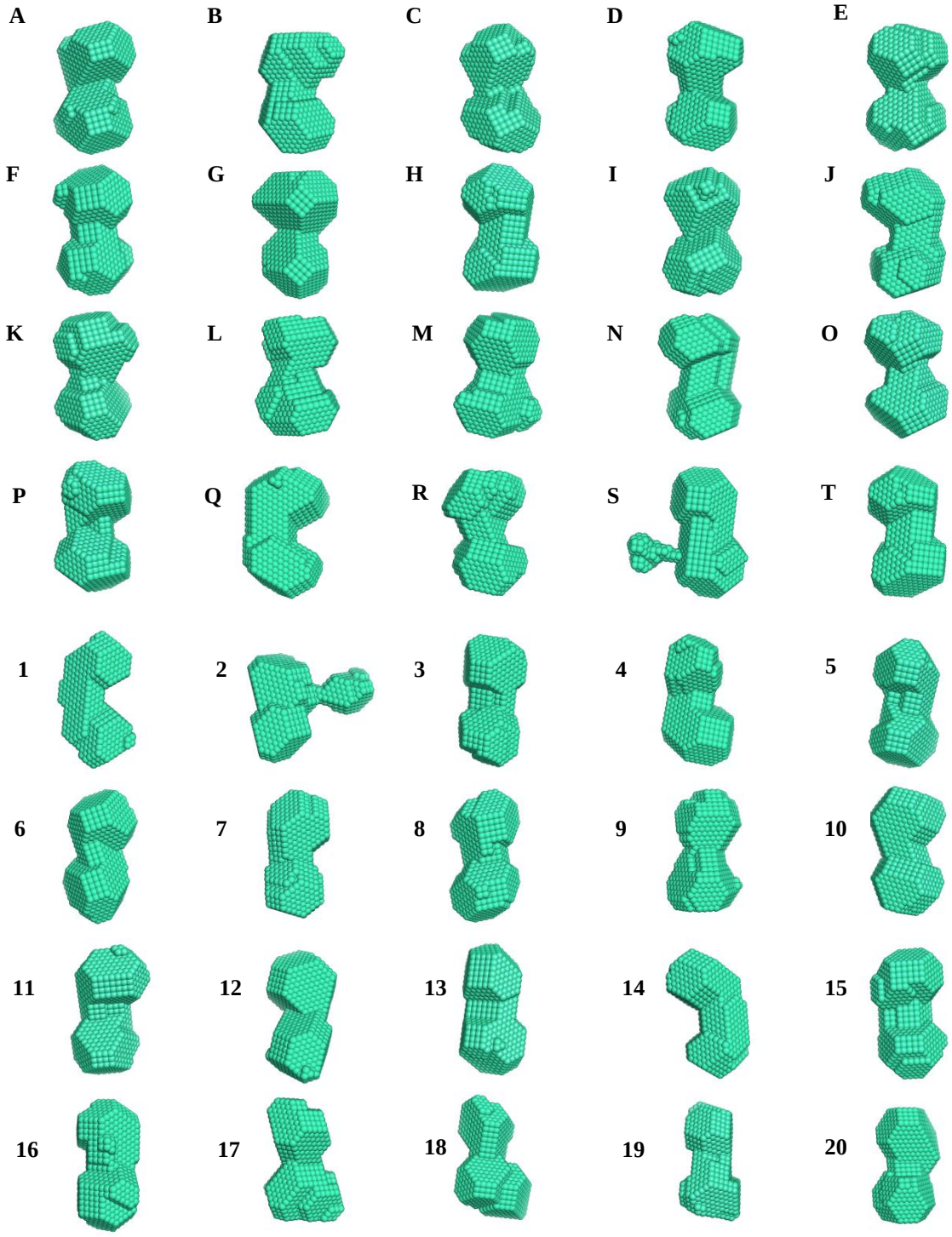


Figure 3.59. Molecular envelopes of untagged BB_0562 generated by DAMMIF. Molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models. Models A to T generated from the 5 mg/ml data and models 1-20 generated from BB_0562 at 3 mg/ml.

3.10.8.6. Evaluation and refinement of untagged BB_0562

The generated untagged envelopes for BB_0562 were evaluated by DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003). Outliers were removed as previously described and the envelopes superimposed on top of the reference set. For the 5 mg/ml untagged data set envelope S was removed and envelopes H and T selected as the reference. For the 3 mg/ml data set envelope 2 was removed and models 7 and 20 used as the reference sets. Full DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003) results are summarised in table 3.17.

Following refinement DAMAVER (Volkov & Svergun, 2003) was used to produce an averaged probability map which was later filtered by DAMFILT (Volkov & Svergun, 2003) to give two final molecular envelopes which are shown in figure 3.60.

Solution	DAMSEL NSD average (Standard deviation of NSD = 0.086)	Include/Discard?	DAMSUP NSD superimposition value	Indicative fit (Chi squared)
	Untagged BB_0562 at 5 mg/ml			
A	0.528	Include	0.452	2.904
B	0.520	Include	0.468	2.900
C	0.533	Include	0.471	2.904
D	0.523	Include	0.517	2.902
E	0.527	Include	0.506	2.904
F	0.523	Include	0.489	2.909
G	0.528	Include	0.514	2.902
H	0.503	Reference	0.000	2.895
I	0.519	Include	0.502	2.913
J	0.528	Include	0.499	2.909
K	0.520	Include	0.510	2.904
L	0.514	Include	0.519	2.900
M	0.530	Include	0.548	2.904
N	0.580	Include	0.563	2.902
O	0.519	Include	0.515	2.903
P	0.534	Include	0.487	2.909
Q	0.556	Include	0.529	2.902
R	0.521	Include	0.508	2.894
S	0.903	Discard	-----	2.912
T	0.504	Reference	0.000	2.911
Solution	**DAMSEL NSD average (Standard deviation of NSD = 0.110)**	**Include/Discard?**	**DAMSUP NSD superimposition value**	**Indicative fit (Chi squared)**
	Untagged BB_0562 at 3 mg/ml			
1	0.585	Include	0.534	4.599
2	1.036	Discard	-----	4.587
3	0.547	Include	0.513	4.597
4	0.546	Include	0.544	4.598
5	0.550	Include	0.496	4.600
6	0.547	Include	0.508	4.588
7	0.516	Reference	0.000	4.606
8	0.543	Include	0.467	4.597
9	0.556	Include	0.496	4.588
10	0.521	Include	0.315	4.599
11	0.567	Include	0.522	4.581
12	0.553	Include	0.530	4.598
13	0.550	Include	0.513	4.597
14	0.564	Include	0.508	4.585
15	0.542	Include	0.496	4.604
16	0.550	Include	0.511	4.612
17	0.549	Include	0.526	4.612
18	0.602	Include	0.638	4.590
19	0.570	Include	0.549	4.592
20	0.516	Reference	0.000	4.597

Table 3.17. A summary of the evaluation and refinement of untagged BB_0562 models by DAMSEL and DAMSUP. DAMSEL (Volkov & Svergun, 2003) determined the mean NSD as 0.546 and 0.575 for the 3 mg/ml and 6 mg/ml data sets respectively. Envelope S from the 3 mg/ml data set and B from the 6 mg/ml data set were calculated as outliers and the data discarded. Data for envelopes H and T from the 3 mg/ml data set and G and T from the 6 mg/ml data set were identified as the most probably solutions for each case and this data used as the reference sets for superimposition. DAMSUP (Volkov & Svergun, 2003) NSD values were then calculated relative to the reference set. The chi squared value represents the indicative fit calculated by DAMMIF (Volkov & Svergun, 2003) for each generated model.

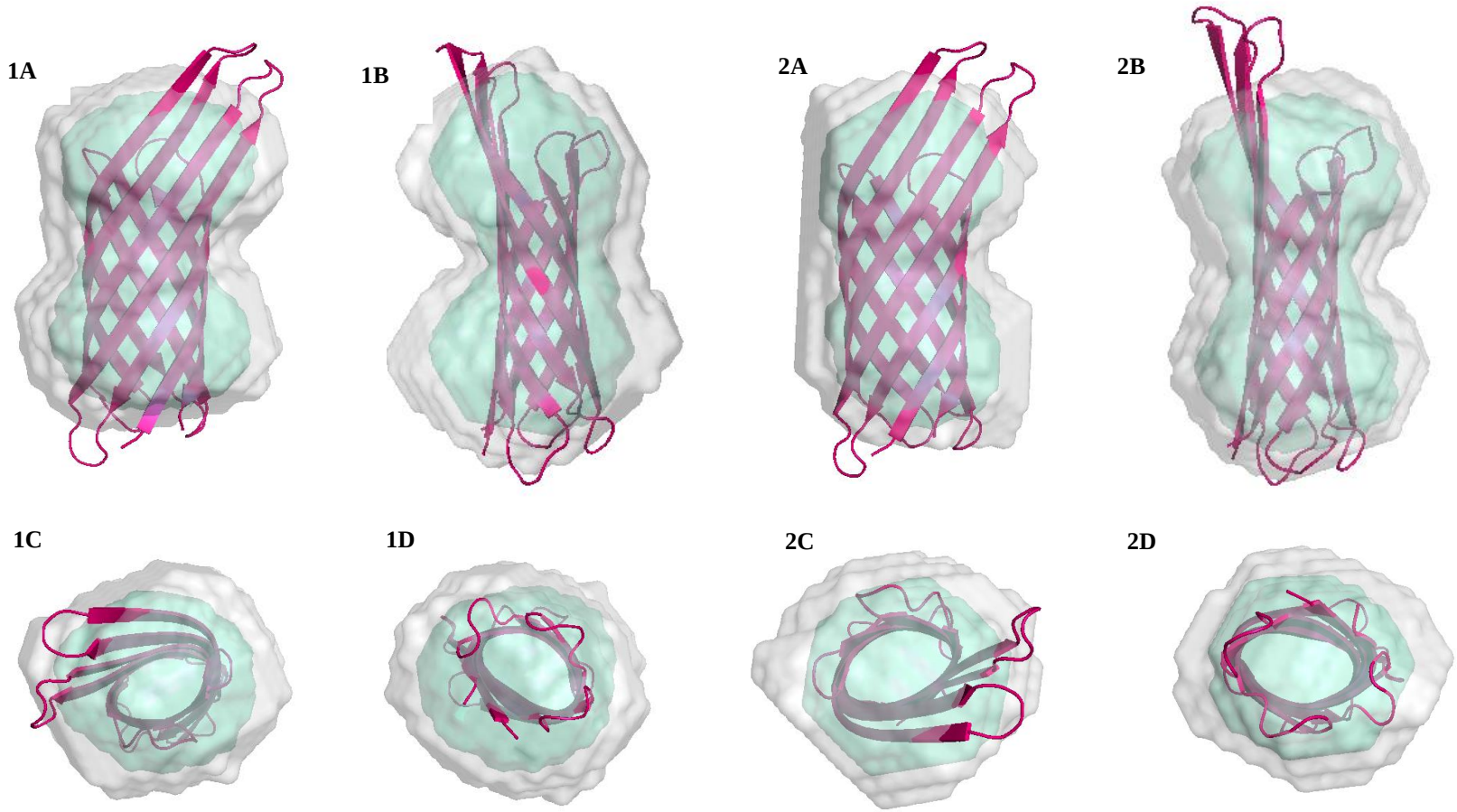


Figure 3.60. Final molecular envelopes for untagged BB_0562. Following the generation of 20 DAMMIF (Franke & Svergun, 2009) models the final envelope was determined after filtering and refinement as previously described. The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope, the grey as the computed probability map generated by DAMAVER (Volkov & Svergun, 2003) and the pink ribbon cartoon is the known crystal structure of *E. coli* OmpX (1QJ8 - Vogt & Schulz, 1999) taken from the protein data bank. All models were generated using Pymol (Schrödinger, 2010). 1A-D) Final molecular envelope of untagged BB_0562 at 5 mg/ml. 2A-D) The final model of untagged BB_0562 at 3 mg/ml.

3.10.8.7. Generating molecular envelopes for poly-histidine tagged BB_0562

Using the same methodology as previously described the background reduced scattering data for tagged BB_05622 was transformed using GNOM (Svergun, 1992) and the output file submitted to DAMMIF (Franke & Svergun, 2009) for ab initio modelling. The process was repeated until twenty impendent molecular envelopes were generated. The initial envelopes are shown in figure 3.61.

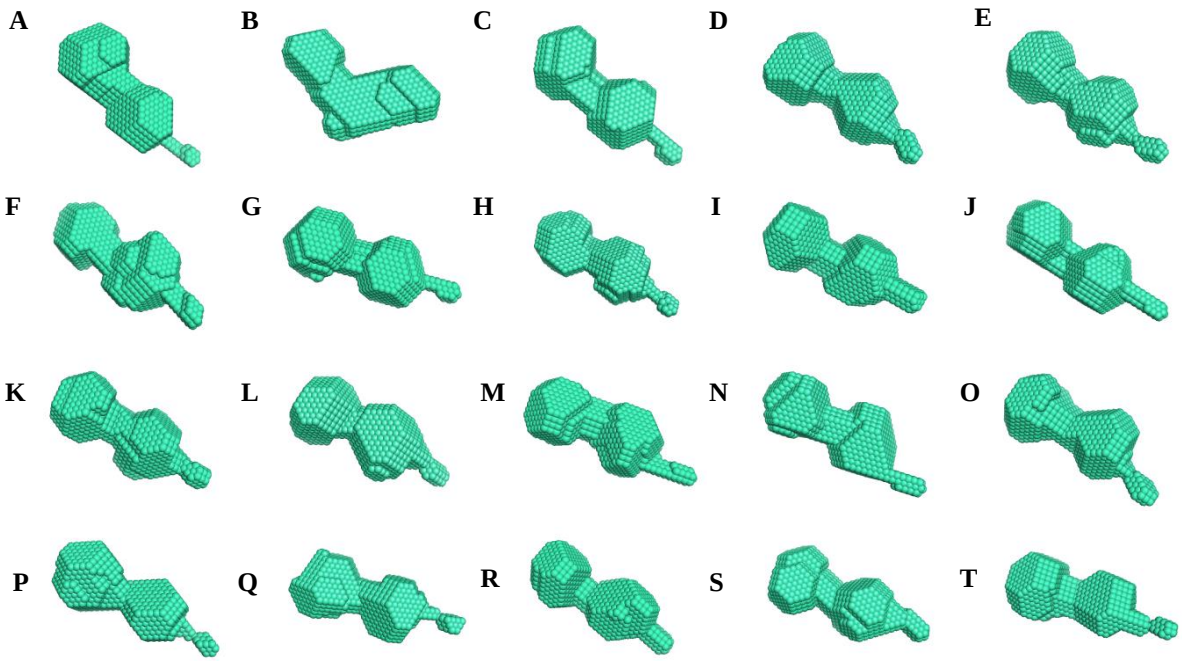


Figure 3.61. Molecular envelopes of histidine tagged BB_0562 at 6 mg/ml. Molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models.

3.10.8.8. Evaluation and refinement of poly-histidine tagged BB_0562

Following the generation of twenty independent models for poly-histidine tagged BB_0562 by DAMMIF (Franke & Svergun, 2009) the PDB files were submitted to DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003) for evaluation. The NSD was calculated for each model and outliers removed based upon a standard deviation of two. Envelope B was deemed as an outlier and removed before superimposition using envelopes D and L as the reference models. Full refinement data is shown in table 3.18.

Following the removal of the outlier and the superimposition of all models the files were used by DAMSEL (Volkov & Svergun, 2003) to produce a probability map which was further filtered by DAMFILT (Volkov & Svergun, 2003) generating two final models as shown in figure 3.62.

Solution	DAMSEL NSD average (Standard deviation of NSD = 0.091)	Include/Discard?	DAMSUP NSD superimposition value	Indicative fit (Chi squared)
Histidine tagged BB_0562				
A	0.544	Include	0.521	4.042
B	0.944	Discard	-----	4.042
C	0.545	Include	0.551	4.040
D	0.516	Reference	0.000	4.044
E	0.548	Include	0.505	4.043
F	0.573	Include	0.576	4.045
G	0.535	Include	0.530	4.040
H	0.534	Include	0.521	4.041
I	0.563	Include	0.532	4.045
J	0.534	Include	0.538	4.042
K	0.549	Include	0.516	4.039
L	0.515	Reference	0.000	4.040
M	0.527	Include	0.506	4.045
N	0.578	Include	0.549	4.037
O	0.549	Include	0.340	4.042
P	0.528	Include	0.535	4.042
Q	0.528	Include	0.535	4.043
R	0.583	Include	0.547	4.042
S	0.554	Include	0.545	4.040
T	0.571	Include	0.534	4.041

Table 3.18. A summary of the evaluation and refinement of histidine tagged BB_0562 models by DAMSEL and DAMSUP. DAMSEL (Volkov & Svergun, 2003) determined the mean NSD as 0.566. Envelope B was calculated as an outlier and the data discarded. Data for envelopes D and L were identified as the most probably solutions in each case and this data used as the reference set for superimposition. DAMSUP (Volkov & Svergun, 2003) NSD values were calculated relative to the reference set. The chi squared value represents the indicative fit calculated by DAMMIF (Volkov & Svergun, 2003) for each generated model.

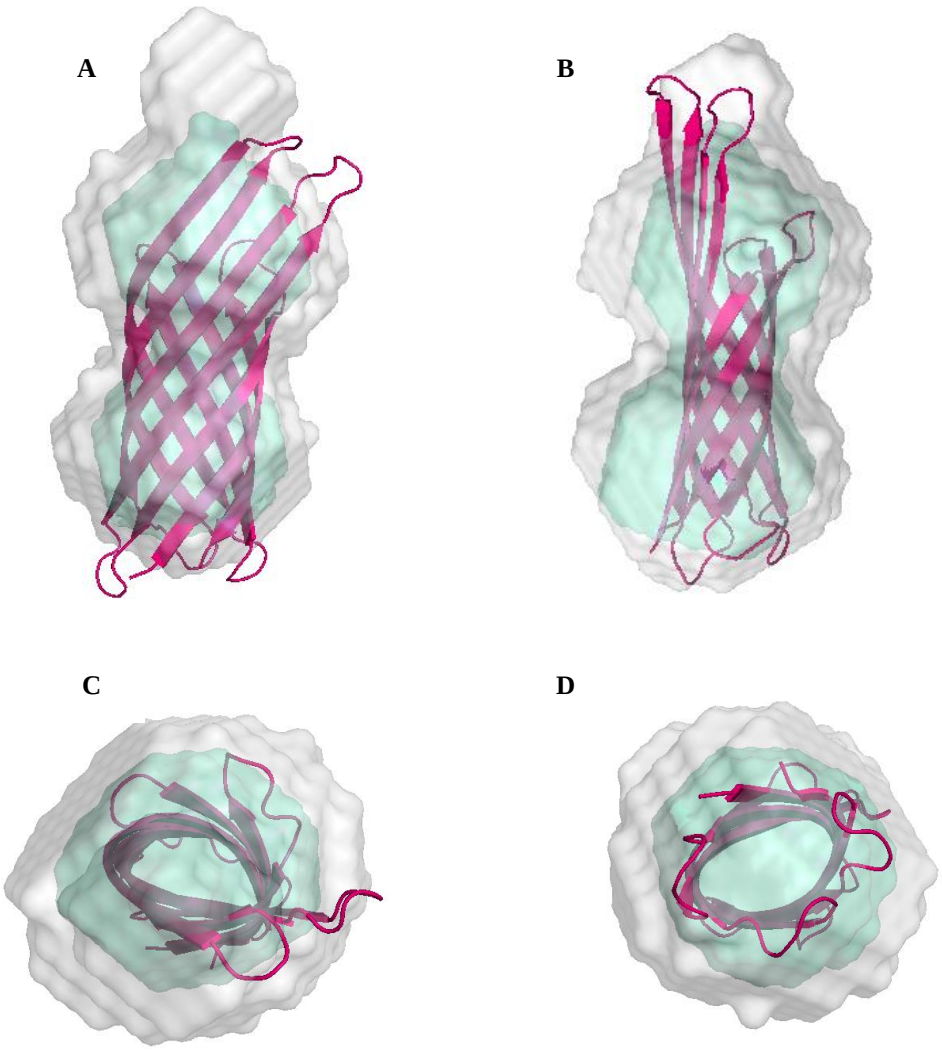


Figure 3.62. Refined and filtered molecular envelope for histidine tagged BB_0562 at 6 mg/ml. Following the generation of 20 DAMMIF (Franke & Svergun, 2009) models the final envelope was determined after filtering and refinement by DAMAVER (Volkov & Svergun, 2003). The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope and the grey is the computed probability map generated by DAMAVER (Volkov & Svergun, 2003). The purple cartoon is OmpX from *E. coli*, data taken from the protein data bank (1QJ8 - Vogt & Schulz, 1999).

3.10.9.1. SAXS data collection for poly-histidine tagged BB_0406

SAXS data was generated for tagged BB_0406 across three concentrations. Data was collected at Diamond Light Source on beamline B21. A total of 18 frames were acquired each lasting three seconds, the resulting scattering was averaged and the blank buffer scattering subtracted resulting in the scattering data shown in figure 3.63.

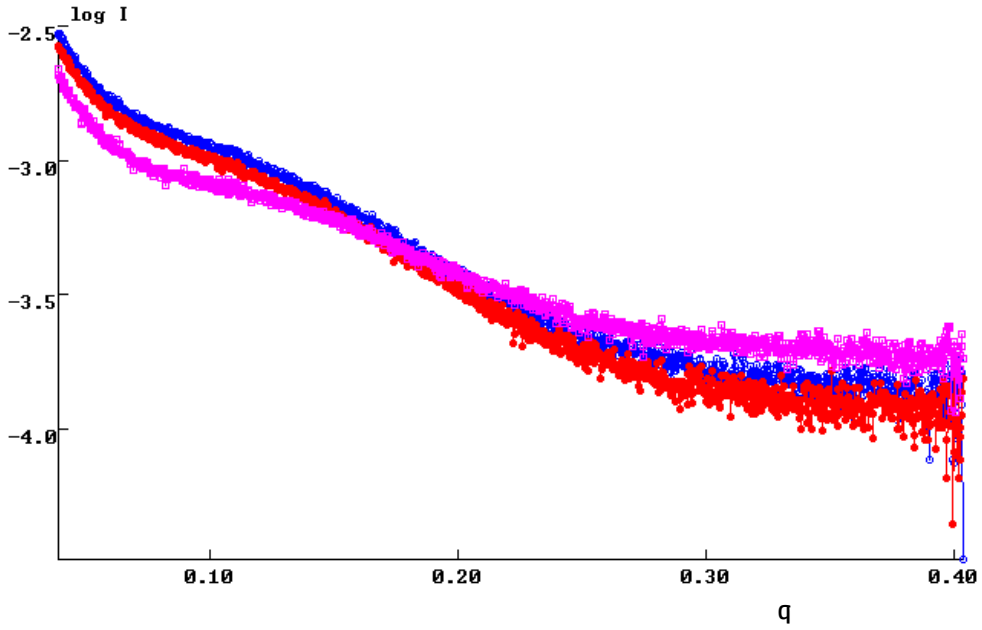


Figure 3.63. Background subtracted scattering data generated for samples of BB_0406 at various concentrations. The background subtracted scattering data for three different concentrations of BB_0406. Recorded scattering data displayed as log of intensity over the scattering vector q (Å). BB_0406 is in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. Background subtraction made with a dialysis matched buffer. Purple data generated from BB_0406 at 3 mg/ml. Blue data generated from BB_0406 at 2 mg/ml. Red data generated from BB_0406 at 1 mg/ml.

All three concentrations of tagged BB_0406 showed an upturn at low q following background subtraction, likely indicative of some level of aggregation within the sample. As preliminary evaluation of the radius of gyration was similar to previous OMP proteins the data was kept and used for modelling. Full Guinier analysis results are summarised in table 3.19. Kratky analysis shown in

figure 3.64 demonstrated that two of the samples (3 and 2 mg/ml) had a defined maxima. However, the curves did not rapidly approach zero. The 1 mg/ml sample was clearly indicating a sample containing partially folded protein and was discarded.

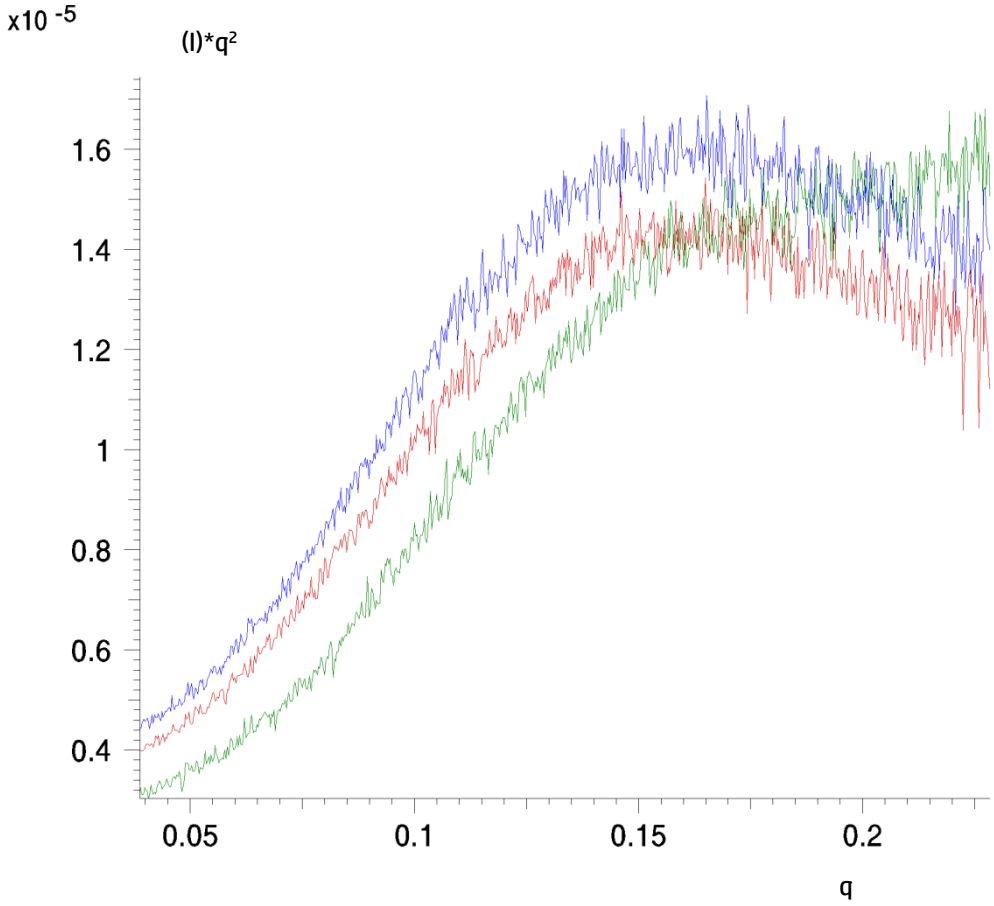


Figure 3.64. Kratky analysis generated from samples of poly-histidine tagged BB_0406. Kratky analysis or $I(q)*q^2$ plot of BB_0406 in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. BLUE – BB_0406 at 3 mg/ml. RED – BB_0406 at 2 mg/ml. GREEN – BB_0406 at 1 mg/ml.

Parameter Calculated	Radius of gyration (Å) 1 mg/ml sample	Radius of gyration (Å) 2 mg/ml sample	Radius of gyration (Å) 3 mg/ml sample
Primus Auto *Rg*	22 +- 0.7 (s*Rg* 1.1-1.5)	15.3 +-0.3 (s*Rg* 0.89-1.3)	15.5 +-0.49 (s*Rg* 0.75-1.235)
Manual Guinier *Rg*	25 +-3 (s*Rg* 0.9-1.6)	15.2 +-0.5 (s*Rg* 0.9-1.33)	14.8 +-0.146 (s*Rg* 0.674-1.50)
Reciprocal space *Rg*	---	14.79	14.90
Real space *Rg*	---	14.8	14.89

Table 3.19. Radius of gyration data for tagged BB_0406. The radius of gyration was calculated for all three concentrations of BB_0406 using several methods. The 1 mg/ml data set was discarded due to the large *Rg* determined from the Guinier analysis and was not taken forward to GNOM (Svergun, 1992) for determination of reciprocal and real *Rg*.

3.10.9.2. GNOM analysis and evaluation of the pair-distance distribution function of poly-histidine tagged BB_0406

Following background subtraction and the initial analysis of the scattering data for BB_0406 the two remaining data sets were transformed using GNOM (Svergun, 1992), the one dimensional scattering curves were taken and the P(r) function was determined and are shown in figure 3.65. The radius of gyration for both samples of BB_0406 at 2 and 3 mg/ml were calculated in both real and reciprocal space and were in good agreement as shown in table 3.18. The P(r) function also was indicative of a compact single domain structure with a selected dmax of 42 Å. GNOM (Svergun, 1992) was also used to produce a .out file in preparation for ab initio modelling later used to generate initial models be DAMMIF (Franke & Svergun, 2009).

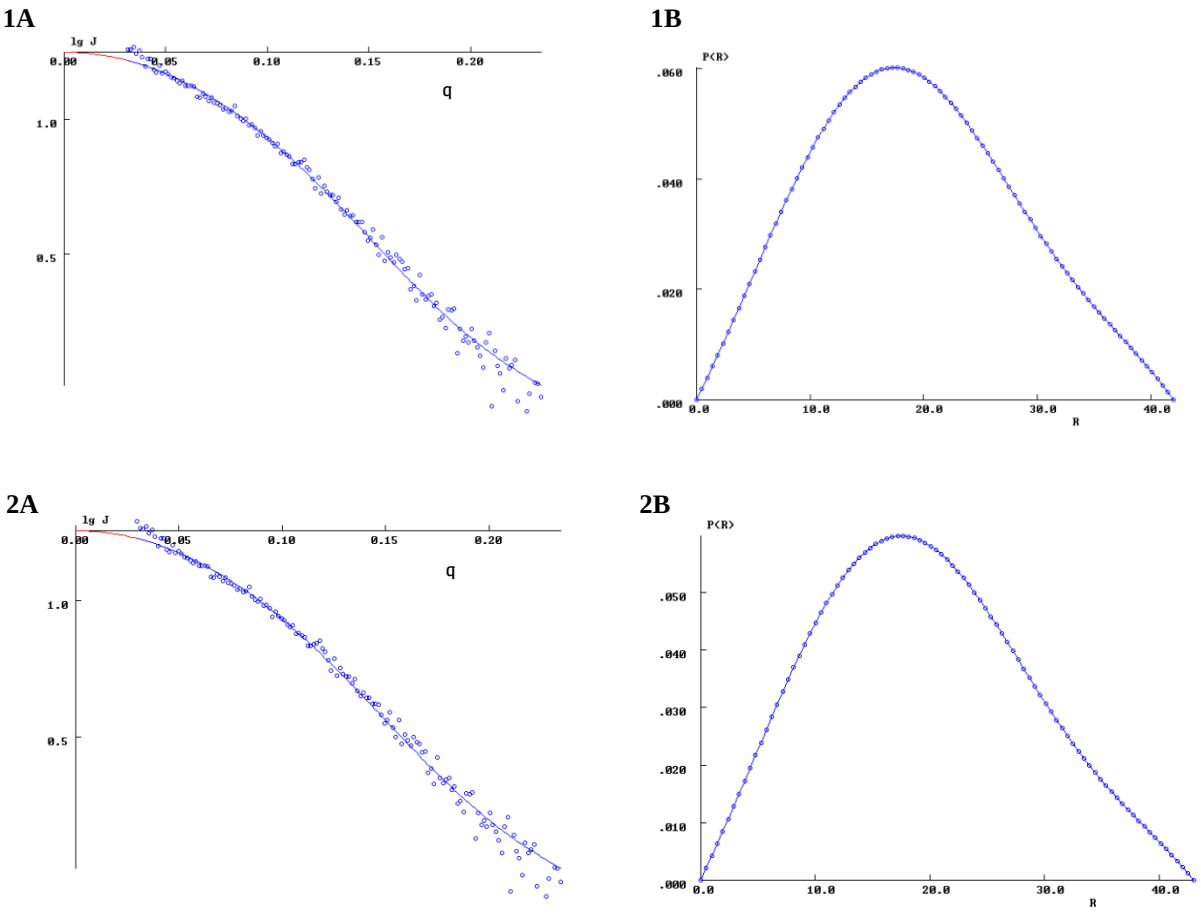


Figure 3.65. GNOM analysis of included BB_0406 data sets. GNOM (Svergun, 1992) was used as an indirect transform tool, the radius of gyration was calculated for all data sets in both real and reciprocal space. **1A** – Reciprocal space fit for tagged BB_0406 at 3 mg/ml. **1B** – P(r) function for tagged BB_0406 at 3 mg/ml. **2A** - Reciprocal space fit for tagged BB_0406 at 2 mg/ml. **2B** - P(r) function for tagged BB_0406 at 2 mg/ml.

3.10.9.3. Generating molecular envelopes for poly-histidine tagged BB_0406

Using the same methodology as previously described, the background subtracted scattering data was transformed using GNOM (Svergun, 1992) and the output file submitted to DAMMIF (Franke & Svergun, 2009) for *ab initio* modelling. The process was repeated until twenty impendent molecular envelopes were produced for each data set resulting in a total of 40 models as shown in figure 3.66.

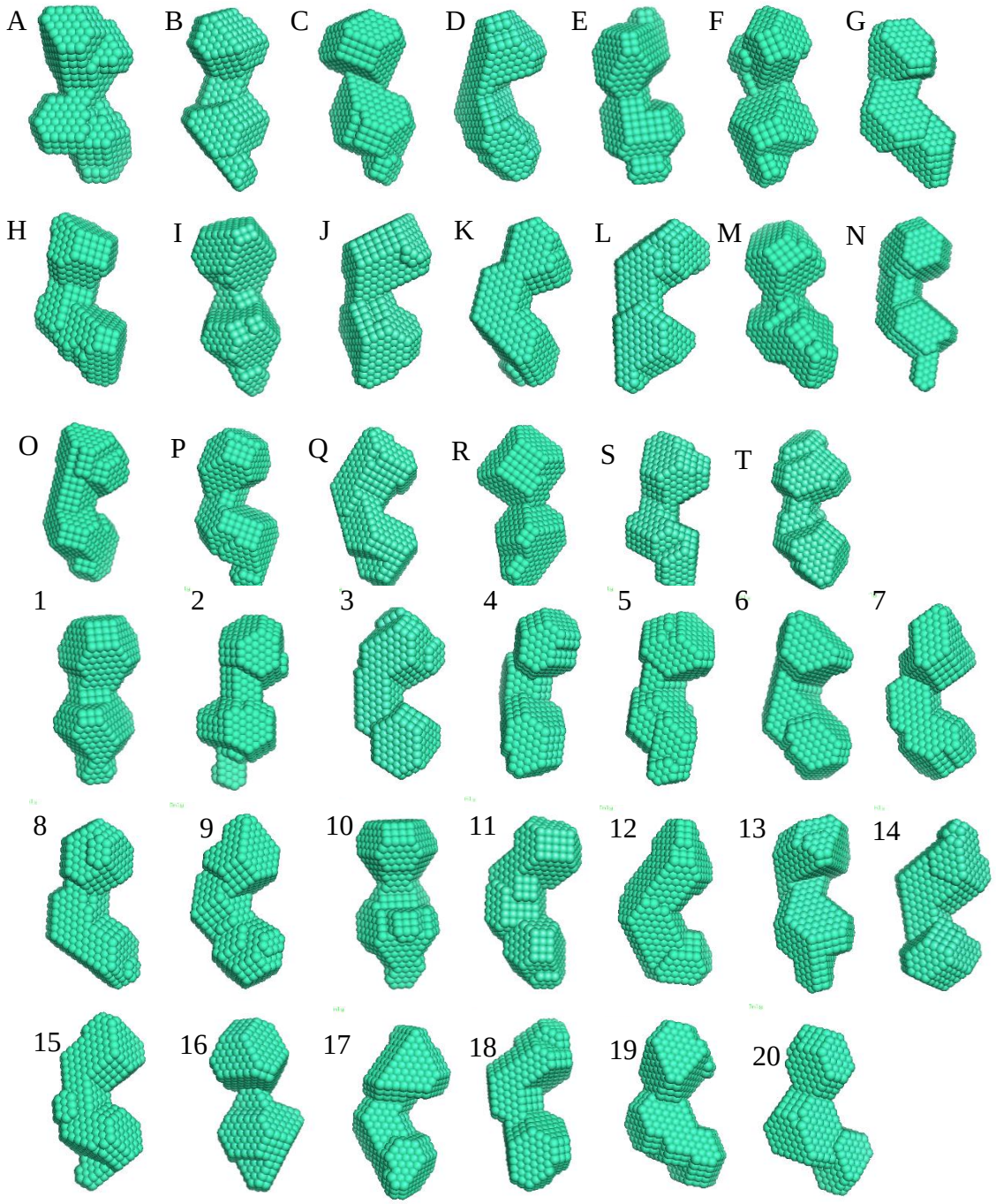


Figure 3.66. Molecular envelopes of poly-histidine tagged BB_0406. Molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models for each data set. Models A-T generated from tagged BB_0406 at 3 mg/ml. Models 1-20 generated from tagged BB_0406 at 2 mg/ml.

3.10.9.4. Evaluation and refinement of poly-histidine tagged BB_0406

Assessment of the generated envelopes was carried out by DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003) as previously described with full refinement results shown in table 3.20. Only one envelope was identified as an outlier across both BG_0408 datasets, envelope K from the 3 mg/ml data set gave a NSD on 0.592 which was beyond two standard deviation points and was removed prior to DAMSUP (Volkov & Svergun, 2003) analysis. Envelopes K and 18 were selected as the reference models and used for superimposition of the remaining solutions.

Following the removal of the outlier and superimposition DAMAVER (Volkov & Svergun, 2003) was used to produce an averaged probability map which was filtered by DAMFILT (Volkov & Svergun, 2003) to give a final molecular envelope for tagged BG_0408 which is shown in figure 3.67.

Solution	DAMSEL NSD average (Standard deviation of NSD = 0.015)	Include/Discard?	DAMSUP NSD superimposition value	Indicative fit (Chi squared)
Tagged BB_0406 at 3 mg/ml				
A	0.552	**Include**	0.525	10.503
B	0.561	**Include**	0.515	10.524
C	0.567	**Include**	0.544	10.566
D	0.555	**Include**	0.531	10.509
E	0.560	**Include**	0.536	10.513
F	0.559	**Include**	0.568	10.510
G	0.559	**Include**	0.531	10.522
H	0.552	**Include**	0.527	10.519
I	0.567	**Include**	0.581	10.377
J	0.535	**Include**	0.539	10.483
K	0.592	**Discard**	-----	10.503
L	0.546	**Include**	0.492	10.471
M	0.555	**Include**	0.517	10.429
N	0.571	**Include**	0.566	10.450
O	0.532	**Reference**	0.000	10.377
P	0.543	**Include**	0.512	10.385
Q	0.552	**Include**	0.554	10.508
R	0.582	**Include**	0.559	10.497
S	0.539	**Include**	0.492	10.510
T	0.541	**Include**	0.470	10.483
Solution	**DAMSEL NSD average (Standard deviation of NSD = 0.027)**	**Include/Discard?**	**DAMSUP NSD superimposition value**	**Indicative fit (Chi squared)**
Tagged BB_0406 at 2 mg/ml				
1	0.612	**Include**	0.624	19.844
2	0.622	**Include**	0.630	19.906
3	0.539	**Include**	0.375	19.902
4	0.562	**Include**	0.488	19.929
5	0.576	**Include**	0.546	19.807
6	0.555	**Include**	0.494	19.992
7	0.549	**Include**	0.520	19.985
8	0.554	**Include**	0.487	19.898
9	0.537	**Include**	0.476	19.887
10	0.603	**Include**	0.611	19.905
11	0.584	**Include**	0.502	19.846
12	0.561	**Include**	0.512	19.881
13	0.587	**Include**	0.610	19.833
14	0.543	**Include**	0.512	19.827
15	0.551	**Include**	0.542	19.972
16	0.615	**Include**	0.638	19.948
17	0.584	**Include**	0.504	19.962
18	0.532	**Reference**	0.000	19.828
19	0.562	**Include**	0.510	19.879
20	0.572	**Include**	0.522	19.943

Table 3.20. A summary of the evaluation and refinement of untagged BB_0406 models by DAMSEL and DAMSUP. DAMSEL (Volkov & Svergun, 2003) determined the mean NSD as 0.556 for the 3 mg/ml sample and 0.568 for the 2 mg/ml sample….. Envelope K was the only model selected as an outlier across both data sets with the data removed before analysis by DAMSUP (Volkov & Svergun, 2003). Envelopes O and 18 were identified as the most probably solutions for the 3 mg/ml and 2 mg/ml data respectively and used as the reference sets for superimposition by DAMSUP (Volkov & Svergun, 2003). The chi squared value represents the indicative fit calculated by DAMMIF (Volkov & Svergun, 2003) for each generated model.

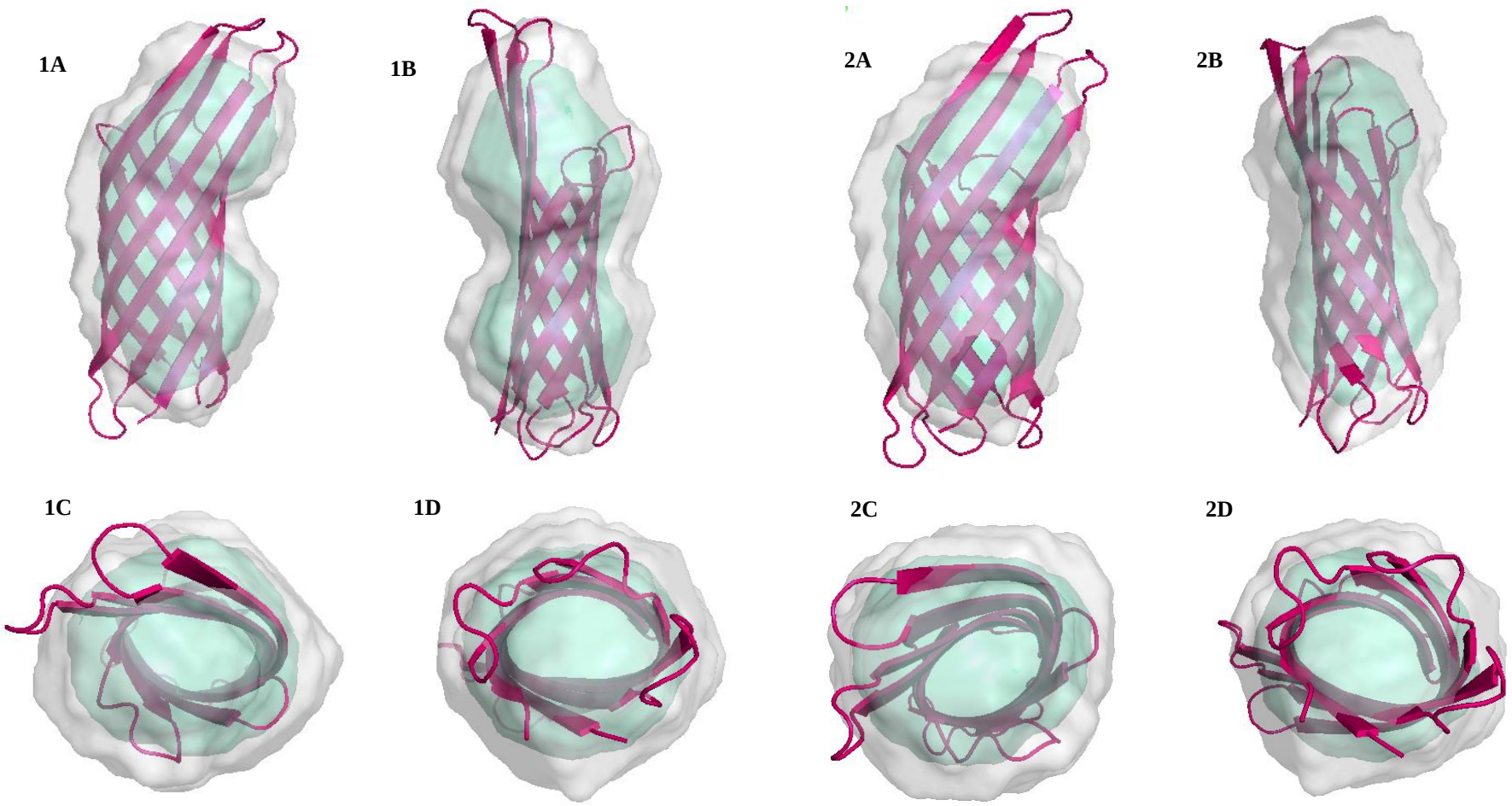


Figure 3.67. Refined and filtered molecular envelopes for poly-histidine tagged BB_0406. Following the generation of 20 DAMMIF (Franke & Svergun, 2009) models for each data set the final envelopes were determined after filtering and refinement by DAMAVER (Volkov & Svergun, 2003). The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope and the grey is the computed probability map generated by DAMAVER (Volkov & Svergun, 2003). The purple cartoon is OmpX from *E. coli*, data taken from the protein data bank (1QJ8 - Vogt & Schulz, 1999). Models 1A-D generated from BB_0406 at 3 mg/ml and models 2A-D generated from BB_0406 at 2 mg/ml.

3.10.10.1 SAXS data collection for BG_0408

SAXS data was generated for untagged BG_0408 at a single concentration of 3 mg/ml and the background subtracted scattering data is shown in figure 3.68. Data was collected for 2400 seconds per frame with ten frames for the sample and blank. Guinier analysis was applied both automatically and manually and is summarised in table 3.21.

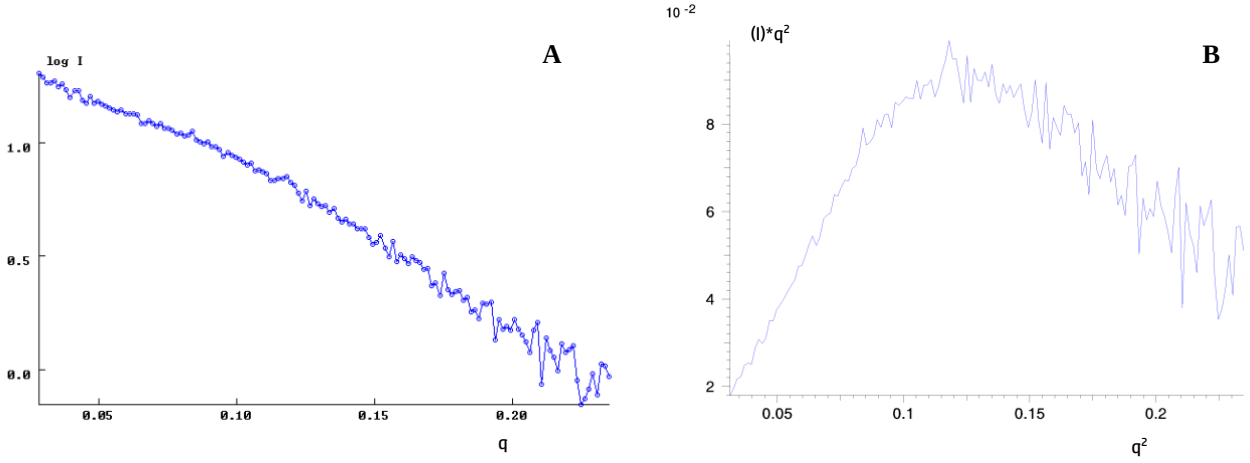


Figure 3.68. Background subtracted scattering data and Kratky analysis generated from a sample of BG0408 at 3 mg/ml. A) The background subtracted scattering data for BG_0408 at 3 mg/ml. Recorded scattering data displayed as log of intensity over the scattering vector q (Å). BG_0408 in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO, pH8. Background subtraction made with a dialysis matched buffer. B) Kratky analysis or I(q)*q^2 plot of BG_0408 at 3 mg/ml. The Kratky plot shows a Gaussian-like bell shape and is indicative of a folded molecule.

Parameter Calculated	***Rg***
Primus Auto *Rg*	15.5 +-0.49 (s*Rg* 0.75-1.235)
Manual Guinier *Rg*	14.8 +-0.146 (s*Rg* 0.674-1.50)
Reciprocal space *Rg*	14.90
Real space *Rg*	14.89

Table 3.21. A summary of the calculated *Rg* for BG0408 at 3 mg/ml. The *Rg* was calculated throughout analysis of the scattering data. The automatically generated *Rg* using Primus (Konarev et al., 2003) appears to deviate from the other values possibly due to poor selection of the Guinier region.

3.10.10.2. GNOM analysis and evaluation of the pair-distance distribution function of BG0408

After the initial analysis of the raw SAXS data, GNOM (Svergun, 1992) from the ATSAS (Svergun, 1992, Franke & Svergun, 2009, Volkov & Svergun, 2003, Kozin & Svergun, 2001, Svergun, 1999) software package is often utilised as an indirect transform program which takes one-dimensional scattering curves and calculates a distance distribution function or P(r) it also prepares files for Ab initio modelling. The P(r) function and reciprocal space fit for the three included data sets are shown in figure 3.69.

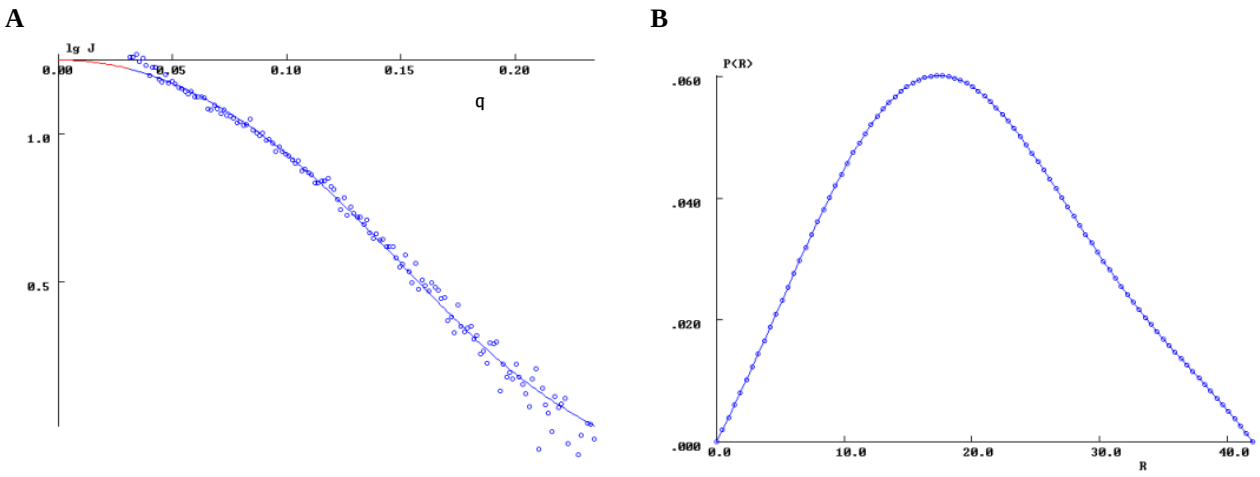


Figure 3.69. GNOM analysis of poly-histidine tagged BG0408. GNOM (Svergun, 1992) was used as an indirect transform tool, the radius of gyration was calculated in both real and reciprocal space. **A** – Reciprocal space fit for tagged BG0408 at 3 mg/ml. **B** – P(r) function for tagged BG0408 at 3 mg/ml.

3.10.10.3. Generating molecular envelopes for poly-histidine tagged BG0408

Using the same methodology as previously described, the raw scattering data was transformed using GNOM (Svergun, 1992) and the output file submitted to DAMMIF (Franke & Svergun, 2009) for *ab initio* modelling. The process was repeated until twenty impendent molecular envelopes were produced and these initial envelopes are shown in figure 3.70.

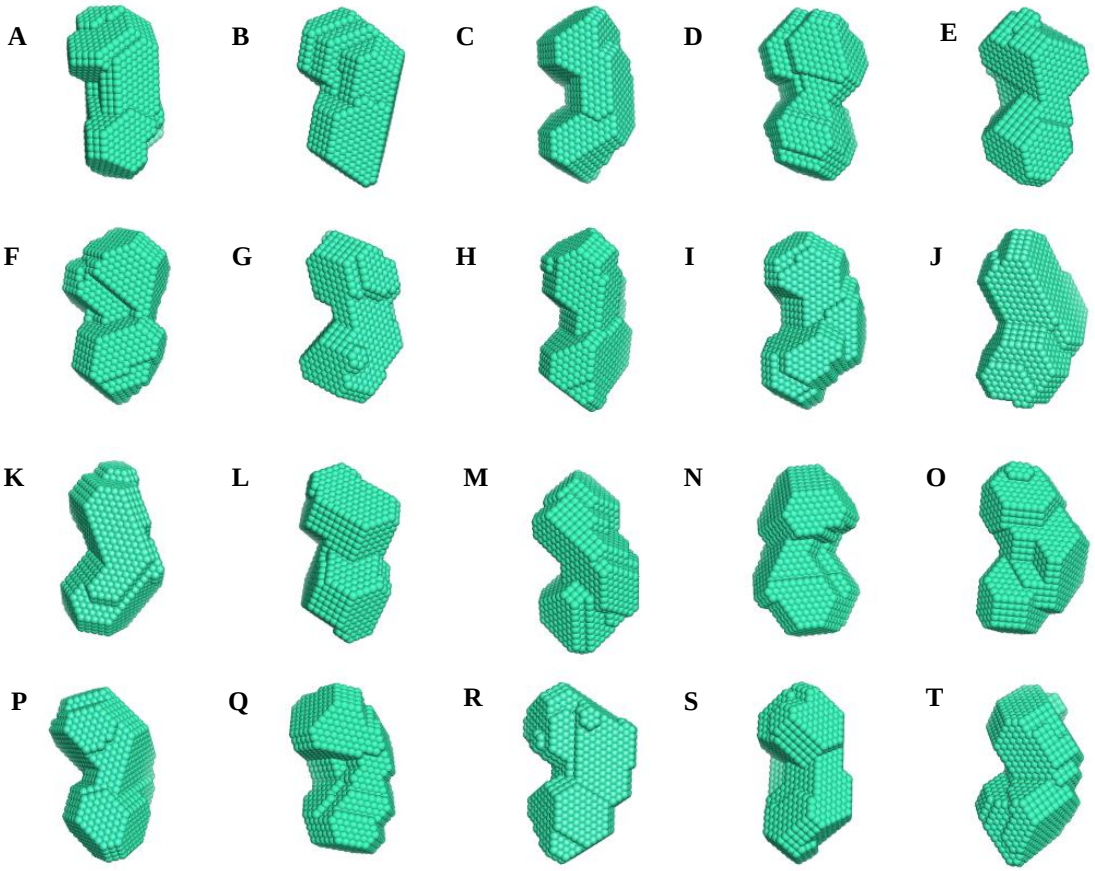


Figure 3.70. Molecular envelopes for poly-histidine tagged BG0408 generated at 3 mg/ml. Initial molecular envelopes were generated by DAMMIF (Franke & Svergun, 2009) and the program repeated to produce twenty independent models.

3.10.10.4. Evaluation and refinement of poly-histidine tagged BG0408

Assessment of the generated envelopes was carried out by DAMSEL (Volkov & Svergun, 2003) and DAMSUP (Volkov & Svergun, 2003) as previously described. No envelopes were deemed as outliers with the mean NSD calculated to be 0.559. Envelope P was selected as the reference model as used by DAMSUP (Volkov & Svergun, 2003) for superimposition with full evaluation data shown in table 3.22.

Following refinement of the DAMMIF (Franke & Svergun, 2009) models DAMAVER (Volkov & Svergun, 2003) was used to produce an averaged probability map which was filtered by DAMFILT (Volkov & Svergun, 2003) to give a final molecular envelope for poly-histidine tagged BG_0408 which is shown in figure 3.71.

Solution	DAMSEL NSD average (Standard deviation of NSD = 0.019)	Include/Discard?	DAMSUP NSD superimposition value	Indicative fit (Chi squared)
Untagged BG_0408				
A	0.593	Include	0.599	1.759
B	0.550	Include	0.530	1.756
C	0.549	Include	0.555	1.757
D	0.577	Include	0.602	1.757
E	0.549	Include	0.521	1.760
F	0.584	Include	0.589	1.758
G	0.534	Include	0.520	1.758
H	0.553	Include	0.536	1.758
I	0.564	Include	0.521	1.757
J	0.594	Include	0.533	1.757
K	0.546	Include	0.319	1.755
L	0.562	Include	0.554	1.757
M	0.561	Include	0.540	1.759
N	0.537	Include	0.495	1.756
O	0.584	Include	0.571	1.760
P	0.532	Reference	0.000	1.757
Q	0.534	Include	0.520	1.755
R	0.550	Include	0.537	1.759
S	0.571	Include	0.514	1.759
T	0.557	Include	0.537	1.757

Table 3.22. A summary of the evaluation and refinement of untagged BG_0408 models by DAMSEL and DAMSUP. DAMSEL (Volkov & Svergun, 2003) determined the mean NSD as 0.559. No outliers were identified based upon the standard deviation of the NSD. Envelope P was identified as the most probably solution and used as the reference set for superimposition and from this DAMSUP (Volkov & Svergun, 2003) NSD values calculated relative to this reference set. The chi squared value represents the indicative fit calculated by DAMMIF (Volkov & Svergun, 2003) for each generated model.

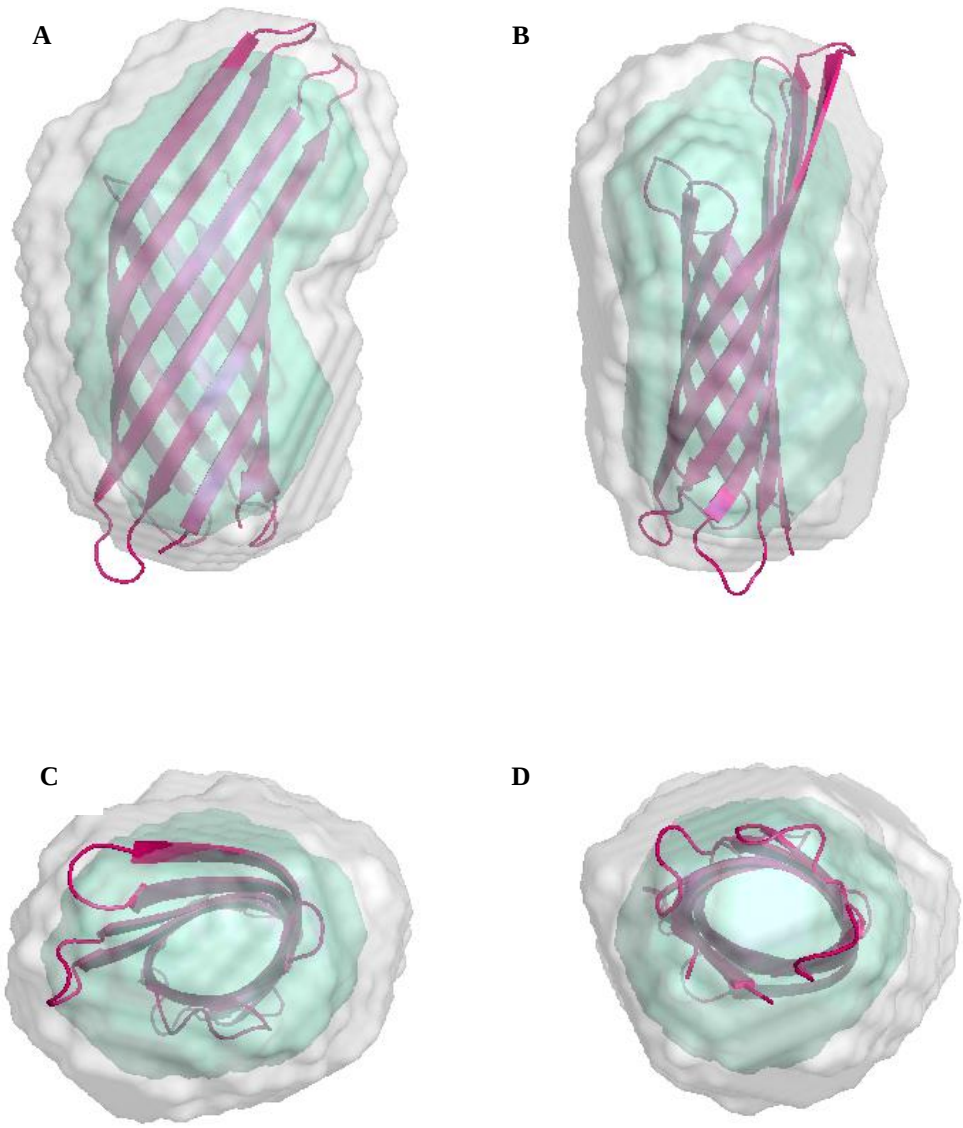


Figure 3.71. Refined and filtered molecular envelope for BG0408 at 3 mg/ml. Following the generation of 20 DAMMIF (Franke & Svergun, 2009) models the final envelope was determined after filtering and refinement by DAMAVER (Volkov & Svergun, 2003). The green represents the DAMFILT (Volkov & Svergun, 2003), filtered envelope and the grey is the computed probability map generated by DAMAVER (Volkov & Svergun, 2003). The purple cartoon is OmpX from *E. coli*, data taken from the protein data bank (1QJ8 - Vogt & Schulz, 1999).

3.10.11 Combining homology modelling and SAXS

All molecular envelopes generated for the four OMP proteins display a distinctive surface shape similar to that of many known 8-stranded β-barrel membrane proteins. Throughout the SAXS analysis the crystal structure of *E. coli* OmpX has been used to show the overall fit of the generated envelope and had been manually docked into the molecular envelope using Pymol (Schrödinger, 2010). However, the four *Borrelia* OMPs may not share the same dimensions or overall shape as the proposed *E. coli* homolog. In order to better understand the fit of each *Borrelia* OMP to the constructed envelope, homology models were generated for all four OMP proteins using Chimera (Pettersen et al., 2004) and Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000) with modelling based upon templates listed in table 2.17 of the methods section. These models were then superimposed into the experimental molecular envelopes as a final assessment of the overall fit with the data shown in figure 3.72. The pdb file coordinates were used to manually dock the homology models into the experimentally determined molecular envelopes. However, software (Colores – Chacón & Wriggers, 2002) is available to perform an exhaustive search of all rigid body degrees of freedom which is useful for multi-domain or multimeric proteins. Following the generation of the homology models for all four *Borrelia* OM proteins the constructed structures were used to compute a theoretical SAXS profile shown in figure 3.73 which was in turn compared to the experimentally determined SAXS profiles. The Chimera (Pettersen et al., 2004) generated homology models for all four *Borrelia* OM proteins were seen to fit well within the experimentally determined molecular envelopes supporting the theory that these possible 8-stranded β-barrel proteins adopt the typical cylindrical shape with similar diameters of known 8-stranded β-barrel crystal structures.

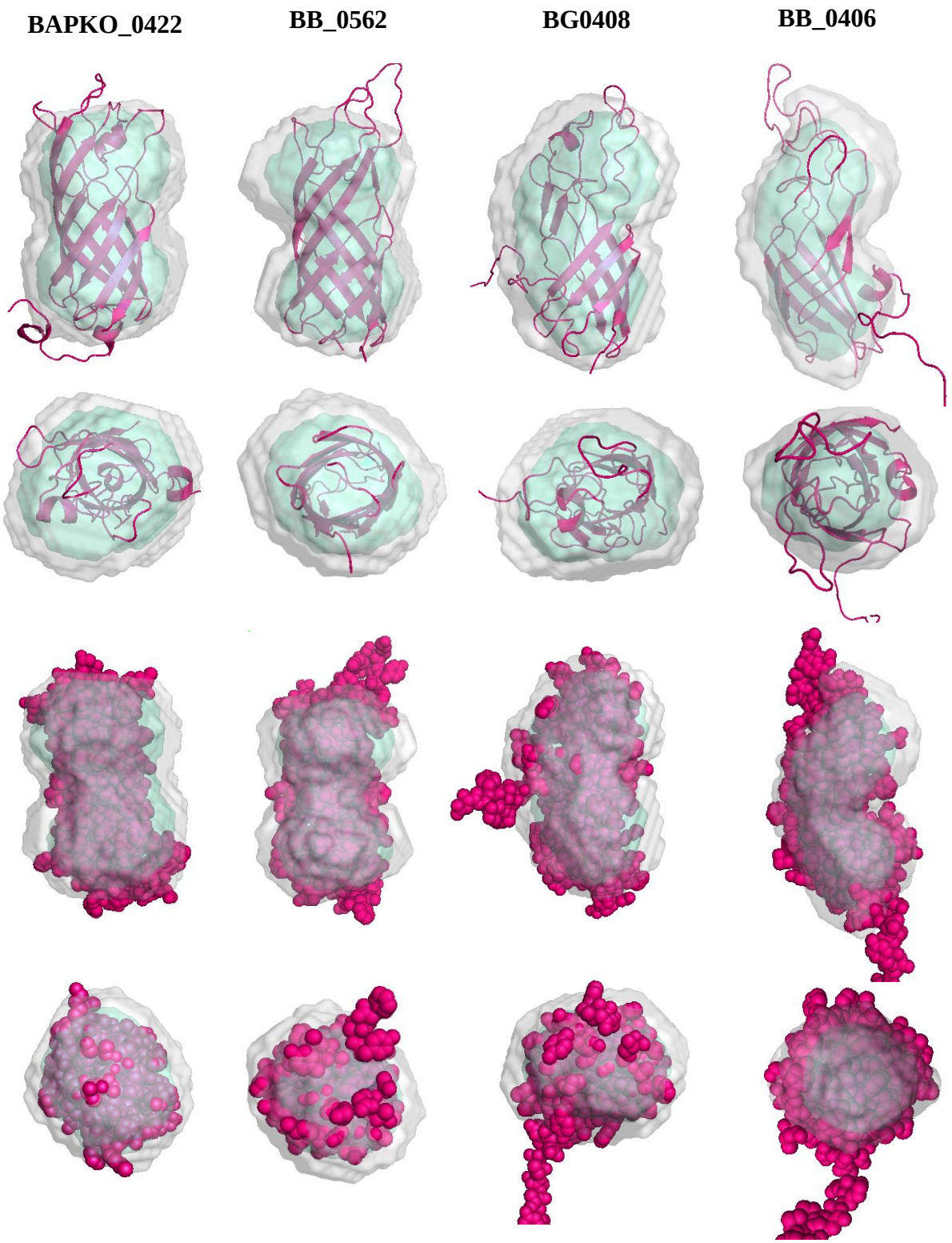


Figure 3.72. Final molecular envelopes for *Borrelia* OMPs with corresponding homology models. The molecular envelopes were produced from the experimental SAXS data with the grey envelope corresponding to the unfiltered DAMAVER (Volkov & Svergun, 2003) model and the green representing the filtered and refined DAMIFILT model. The pink superimposed homology models were created in silico using Chimera (Pettersen et al., 2004) and Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000).

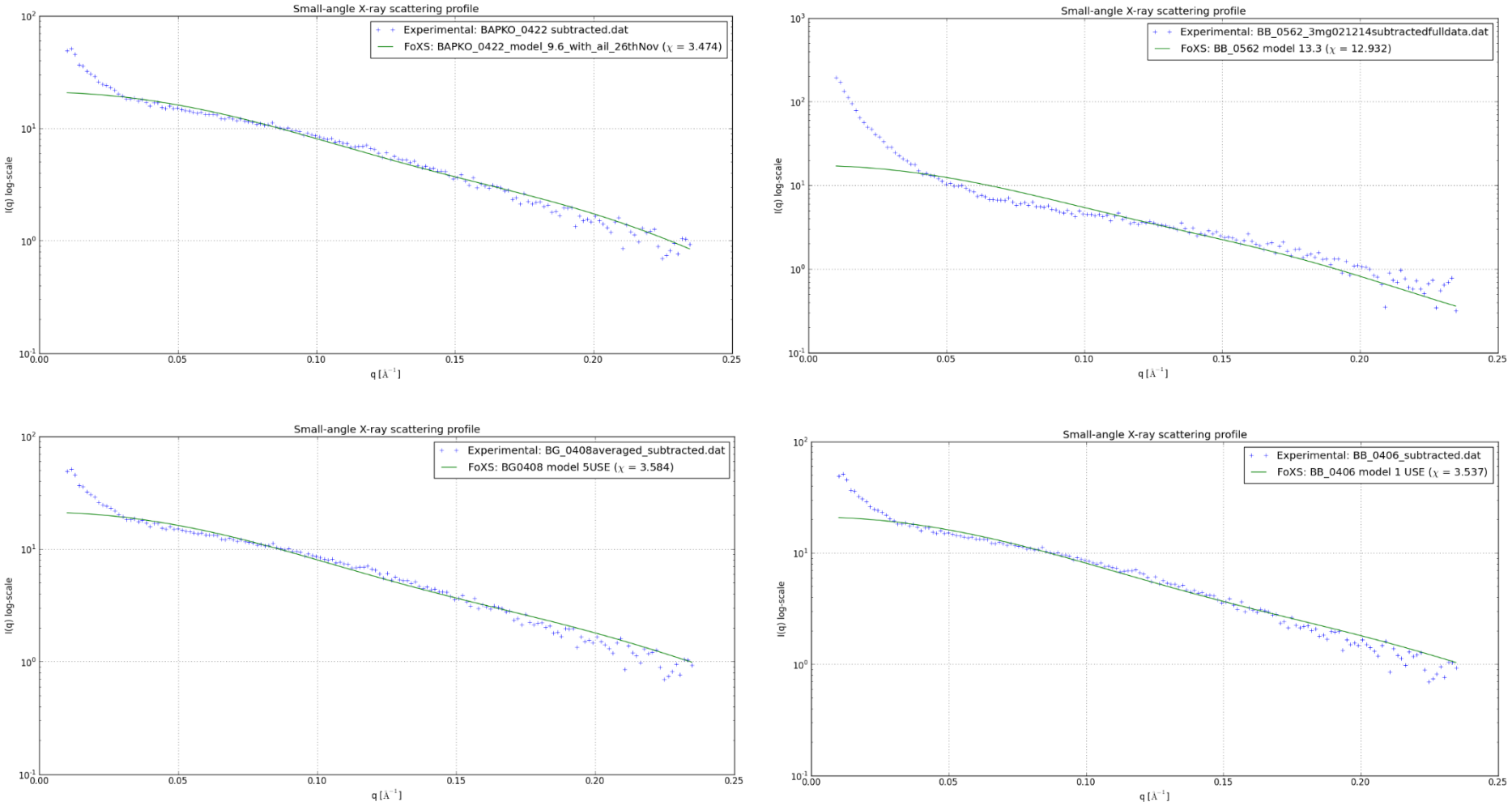


Figure 3.73. Computed theoretical SAXS profiles for *Borrelia* OMP homology models. Theoretical intensity curves generated by Chimera (Pettersen et al., 2004) using the constructed homology models. The theoretical profiles are shown as a green solid line with the experimental SAXS data for each recombinant protein shown as blue crosses. Plots were generated using Chimera (Pettersen et al., 2004).

3.10.12. Overall measurements and dimensions of *Borrelia* OMP molecular envelopes

Measurements for the experimental molecular envelopes and the homology model computed surface were calculated using Pymol (Schrödinger, 2010). Full dimensions are shown in table 3.23. Overall dimensions for all four molecular envelopes for the *Borrelia* OM proteins compare well with the sizes of known 8-stranded β-barrel proteins with examples shown in table 3.24.

Parameters (Å)	**BAPKO_0422 untagged**	**BAPKO_0422 polyhistidine tagged**	**BB_0562 untagged**	**BB_0562 polyhistidine tagged**	**BG0408 polyhistidine tagged**	**BB_0406 polyhistidine tagged**
Radius of gyration	14.38	16.74	14.22	14.31	14.89	14.85
DMAX	42	44	45	45	42	42
Estimated molecular weight (kDa)	20.2	22.9	17.7	20.4	22.8	22.8
Average length of the molecular envelope	45.1-49.2	44.3-50.1	44.6-48.1	44.6-51.8	51.4-58.6	45.3-56.7
Average length of the homology model	57.8	57.8	57.5	57.5	58.9	59.2
Average diameter of the molecular envelope	20.2-26.2	14.8-19	15.7-21.9	15.3-23.7	23.1-33.7	12.9-20.2
Average diameter of the homology model	22.4	22.4	23.5	23.5	23.6	21.2

Table 3.23. Comparing the dimensions of the models for *Borrelia* OM proteins. Estimated molecular weight calculated minus the presumed signal sequence

Protein	PDB entry	PDB reference	Molecular weight (kDa)	Length (Å)	Diameter (Å)
OmpA (*E. coli*)	1BXW	Pautsch and Schulz, 1998	35	61.4	22.3
OmpW (*E. coli*)	2F1T	Hong et al., 2006	21	57.6	22.2
OmpX (*E. coli*)	1QJ8	Vogt & Schulz, 1999	18.6	54.9	24.2
Ail (*Yersinia pestis*)	3QRC	Yamashita et al., 2011	17.5	59.2	19.2
OPRG (*Pseudomonas aeruginosa*)	2X27	Touw et al., 2010	25.2	71.8	22
TT_C0834 (*Thermus thermophillus*)	3DZM	Brosig et al., 2009	23.7	58.6	25.1
NspA (*Neisseria meningitidis*)	1P4T	Vandeputte-Rutten et al., 2003	18.4	52.6	24.5

Table 3.24. Overall sizes and dimensions of known OMP structures. PDB files for each protein were taken from the Protein Data Bank and opened using Pymol (Schrödinger, 2010). Dimensions were calculated using the measurement function in Pymol (Schrödinger, 2010) with calculations based upon the surface exposed model.

3.10.13. SAXS Discussion

Following the production of recombinant protein for the four proposed OmpA-like *Borrelia* proteins, SAXS experiments were designed and carried out with the intention to fulfil the below criteria.

a) To determine a low resolution molecular envelope for each of the four proteins.
b) To compare the low resolution molecular envelopes with surface representations of known 8-stranded β-barrels.
c) To determine the radius of gyration.
d) To calculate the overall dimensions and compare with known 8-stranded β-barrels.
e) To construct homology models and compute a theoretical SAXS profile to be compared with the experimentally obtained data.

Prior to collecting data for the four recombinant proteins the methodology was extensively tested using lysozyme and ribonuclease A as control proteins. The objective of this initial testing was to ensure that the chosen detergent LDAO could be successfully subtracted from the experimental scattering and to explore the issues which surround poor background subtraction. Both lysozyme and ribonuclease A were reconstituted in the same buffer which was used for the four OmpA-like proteins and contained 0.1% v/v LDAO. Data collected and subtracted with the appropriate blank for both proteins produced a molecular envelope in good agreement with models based upon data generated without detergent. This confirmed that for these particular proteins the scattering of the detergent LDAO could be subtracted successfully from the experimental data on the condition that the blank was a good match. When an incorrect blank was subtracted from the same experimental data the resulting envelope resembled a detergent micelle underlining the importance of ensuring that the buffer blank is a true match to the experimental sample.

Although lysozyme and ribonuclease A were extremely useful for testing the proposed methodology the limiting factor was that neither of these are membrane proteins. As membrane proteins contain an exposed hydrophobic region there is an expectation that the detergent LDAO will behave differently within this environment. Prior to processing the experimental data it was expected that the generated molecular envelope may have additional bulking around the proposed hydrophobic girdle. This however did not appear to be the case and detergent subtraction did not give rise to any complications when extra care was taken preparing buffer blanks used for subtraction. These buffer blanks were created by extensive dialysis it was found that a five day dialysis with only one buffer change within the first two hours was sufficient to produce a quality buffer blank for SAXS experiments.

Aggregation observed in SAXS experiments is often hugely detrimental to the resulting data. It is often detected by a sharp upturn at high q or a curvature of the data or residuals in a Guinier analysis. Some levels of aggregation can be alleviated on the condition that the *Rg* is of an expected size and that the upturn in the data does not drastically effect the overall curve. With the exclusion of the control proteins all experimental samples barring one displayed some levels of aggregation as expected with refolded recombinant membrane proteins. Although assessment of aggregation can easily be undertaken by eye, aggregation levels can also be estimated using the auto*Rg* function in Primus (Konarev et al., 2003). Primus (Konarev et al., 2003) estimated the levels of aggregation of 1-3% and as the calculated *Rg* was within the expected range the data was deemed usable. The data also matched well with the sample with no aggregation indicating the small amount of aggregation in the other samples was minimal enough not to effect contrast and cause sufficient changes to the curve.

The two *Borrelia* proteins BB_0406 and BG0408 were particularly problematic for SAXS experiments. Both proteins contain two cysteines eight residues apart at positions 31 and 40. As previously discussed in section 3.4.4. Not only was the final yield of protein poor these proteins also have the ability to appear as monomeric, dimeric, tetrameric and octameric (as shown demonstrated by SDS-PAGE in figure 3.13 of section 3.4.4. The ability of these oligomers to form appears to be dynamic and at some kind of equilibrium making separation difficult (Weber, 1983, Chadwick et al., 1997). Addition of fresh reducing agents including DTT and β-mercaptoethanol at normal concentrations failed to produce a monomeric sample. These difficulties resulted in fewer samples for data collection however one sample was prepared for BG0408 and three samples for BB_0406 by addition of β-mercaptoethanol at 100mM followed by immediate data collection.

Throughout the experiments there was also some discrepancy between the calculated P(r) function and the final DAMMIN (Svergun, 1999) bead model. The P(r) distribution function describes the paired-set of all distances between points within a molecule. The function is also considered as a smooth, non-negative curve when approaching dmax (maximum particle dimensions). As SAXS is a low resolution technique, the measurements are limited by the experimental resolution. Therefore, the P(r) can be thought of as resolution limited sampling of the P(r) distribution function. In most cases the calculated P(r) function was indicative of something between a globule and cylindrical molecule. The P(r) curves (Figures 3.50, 3.58, 3.65, 3.69) are not Gaussian curves which would represent a perfect sphere but instead tail off towards zero which is a feature of cylindrical molecules. In contrast the final bead model was more cylindrical or dumbbell shaped. These differences are due

to the low resolution limits of SAXS. P(r) functions calculated using synchrotron data were seen to be more accurate and represented something between a globular molecule and a cylindrical structure which agreed more closely with the final DAMMIN (Svergun, 1999) bead model. This was due to the higher level of resolution obtained using synchrotron data over in-house SAXS. The dumbbell shape often seen in the final DAMMIN models (Svergun, 1999) is likely an artefact of software averaging. As the protein molecules are positioned in no particular orientation the end model is an average model over three dimensions. As the loop regions at each end of the Borrelia OMPs are considered the most flexible components they may be considered as the most averaged portion leading to a rounding or dumbbell shape.

The dmax or the maximum particle dimension was similar for all *Borrelia* OMPs and was between 42-45 Å. When the dmax was selected within this range the P(r) function was smooth and non-negative at dmax. This estimated dmax value was slightly lower than the average length of the experimentally generated molecular envelopes which ranged from 44.3-51.4 Å for the filtered model (full dimensions and parameters can be reviewed in table 3.22). Homology models were also slightly larger with the average length of OMPs at ~57 Å. The differences observed can be explained by the way each measurement was deduced. The dmax is an estimate of the maximum protein size prior to bead modelling. The DAMFILT (Volkov & Svergun, 2003) envelopes which are the final refined and filtered models agreed well with dmax. The DAMAVER (Volkov & Svergun, 2003) envelopes were larger than dmax yet this is expected as this is the unrefined and unfiltered model. Measurements taken from homology models made *in silico* are theoretical and have not been experimentally obtained. It is likely that the OMPs are more compact than the calculated homology model which would explain the increase seen within the average length of the model when compared to the SAXS envelope.

The *Borrelia* OMP homology models were used to generate theoretical SAXS profiles (figure 3.73) using FoXS (Schneidman-Duhovny et al., 2010). In order to assess the level of agreement between the experimental SAXS curve and the theoretical curve derived from the homology models the chi squared fit was calculated. The chi-square value represents the overall agreement between the two data sets by measuring the error-weighted score (Rambo & Tainer, 2013). However, this function can be unreliable with noisy datasets (Rambo & Tainer, 2013). The chi-squared values are shown in figure 3.73 and were reasonable at ~3.5 for BAPKO_0422, BB_0406 and BG0408. For the case of BB_0562 the chi-squared value was 12.9 indicating the agreement between the two curves was lower than that

observed for the other *Borrelia* OMPs. This higher chi-squared value is likely due to low levels of aggregation within the sample.

In summary the molecular envelopes generated for the four *Borrelia* proteins are consistent with a small transmembrane β-barrel structure, the diameter and overall length of the envelopes are close to literature values for known 8-stranded β-barrels and generated homology models fit well within the generated surface structure.

3.11. Crystallisation attempts of *B. burgdorferi s.l.* OmpA-like proteins

3.11.1. Vapour diffusion crystallisation of BAPKO_0422 and BB_0562

Attempts were made to crystallise recombinant BAPKO_0422 and BB_0562 both using traditional methods and lipidic cubic phase crystallisation. The remaining *Borrelia* proteins BB_0406 and BG0408 were excluded from these trials due to the difficulty in obtaining enough high quality sample. As described in section 2.2.9, BAPKO_0422 and BB_0562 were refolded, purified and concentrated prior to crystallisation attempts. Where possible protein samples were fresh and subjected to refrigerated centrifugation and the concentration assessed using a Jenway Genova Nano, Micro-volume spectrophotometer prior to being placed into crystallisation trays. Hanging drop trays were set up for both BAPKO_0422 and BB_0562 using concentrations ranging from 6-14 mg/ml using seven crystallisation screens described in section 2.2.9 and full formulations can be found within the appendix. Hanging drop trays were stored at both 10 and 20^{o}C and checked at regular intervals. Many conditions resulted in rapid precipitation often within the first couple of days for these cases optimisation was attempted where the precipitant was reduced in concentration however this did not result in any protein crystals. Some conditions did give rise to crystal growth, these were however confirmed to be salt crystals.

3.11.2 Lipidic cubic phase crystallisation of BAPKO_0422

Following the lack of success using traditional crystallisation methods lipidic cubic phase crystallisation was attempted for BAPKO_0422. Lipidic cubic phase (LCP) or *In meso* crystallography exploits the fact that membrane proteins are most stable within lipid bilayers and uses a complex mixture of hydrated lipid often monoolein, protein in solution, detergent and precipitants in order to support crystal growth (Aherne et al., 2012). The lipidic cubic phase is composed of a single lipid bilayer which creates an infinite periodic minimal surface (IPMS) which acts to divide space into two networks of non-intersecting water channels (Landau & Rosenbusch, 1996). Bicontinuous lipidic cubic phases can be separated into three types based upon their space group symmetry these phases are known as Ia3d, 1m3m and Pn3m which are summarised in figure 3.75.

Although LCP's have been used for many applications from drug delivery to biosensors these phases were found particularly useful for the crystallisation membrane proteins (Landau & Rosenbusch, 1996, Srivastava et al., 2014, Tiefenbrunn et al., 2011, Liao et al., 2012, Rasmussen et al., 2011). The first application of LCP for the crystallisation of a membrane protein was achieved in 1996 by Landau

& Rosenbusch. The attraction of LCP crystallisation is that it recognises that often membrane proteins are manipulated in harsh detergent environments as opposed to the more native lipid environment LCP provides.

It has also been suggested that crystals obtained using LCP have better crystal ordering this was proposed as crystals grown in this manner have type 1 packing with protein molecules forming contacts through both hydrophilic and hydrophobic portions reducing the solvent content.

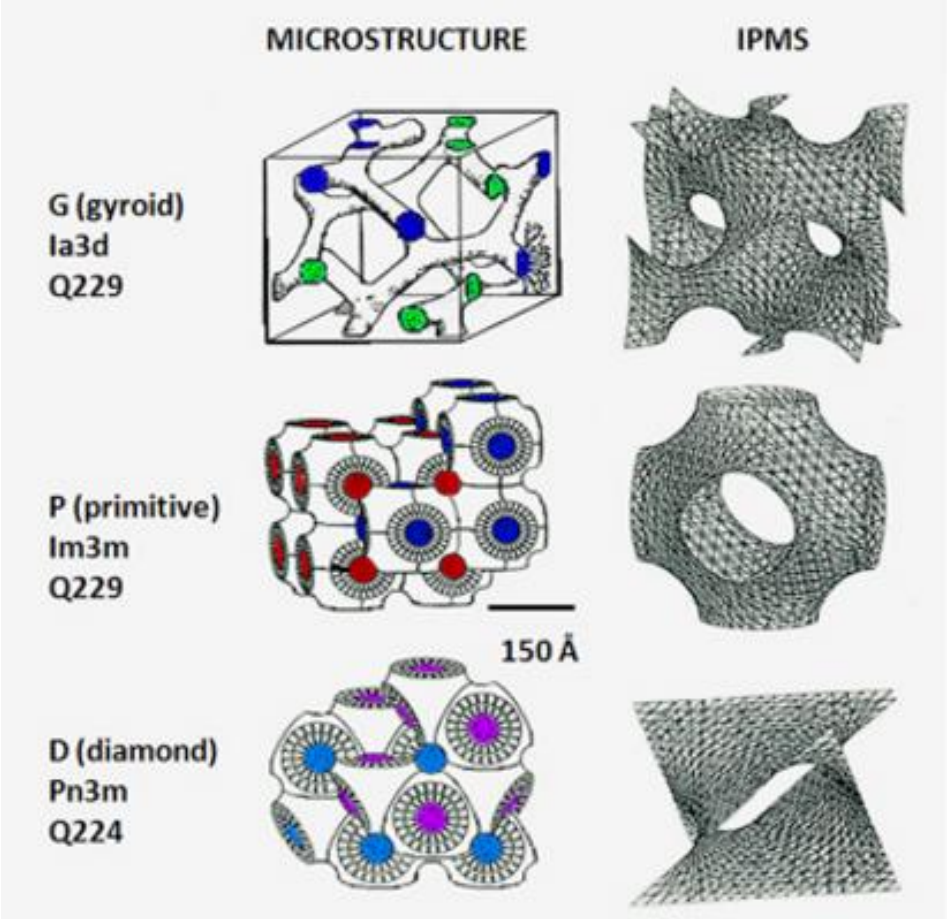


Figure 3.74. The three bicontinuous LCP's and their space groups. The two non-intersecting water channels are represented in two different colours. Figure taken from Qiu & Caffrey, 2000.

Although several lipids can be used for LCP crystallisation monoolein is most commonly the lipid of choice. When mixed with water it spontaneously swells to form distinctive mesophases as further discussed in figure 3.74. At room temperature and at a water concentration of around 40% the *Pn3m* cubic phase is formed, the water concentration is reduced and the phase moves into the further cubic phase *la3d.* Either of these phases can be used when beginning in meso crystallisation but further dehydration after experimental set up is believed to be the key driving force of crystallisation as the

cubic phase moves to a lamellar phase as further discussed in figure 3.75. It is however important to ensure the protein solution is incorporated into the cubic phase and to understand the changes in phase behaviour for the supporting detergent, as the cubic phase is easily disrupted by the addition of detergents and precipitants inducing transition from the cubic to the lamellar phase during experimental set up.

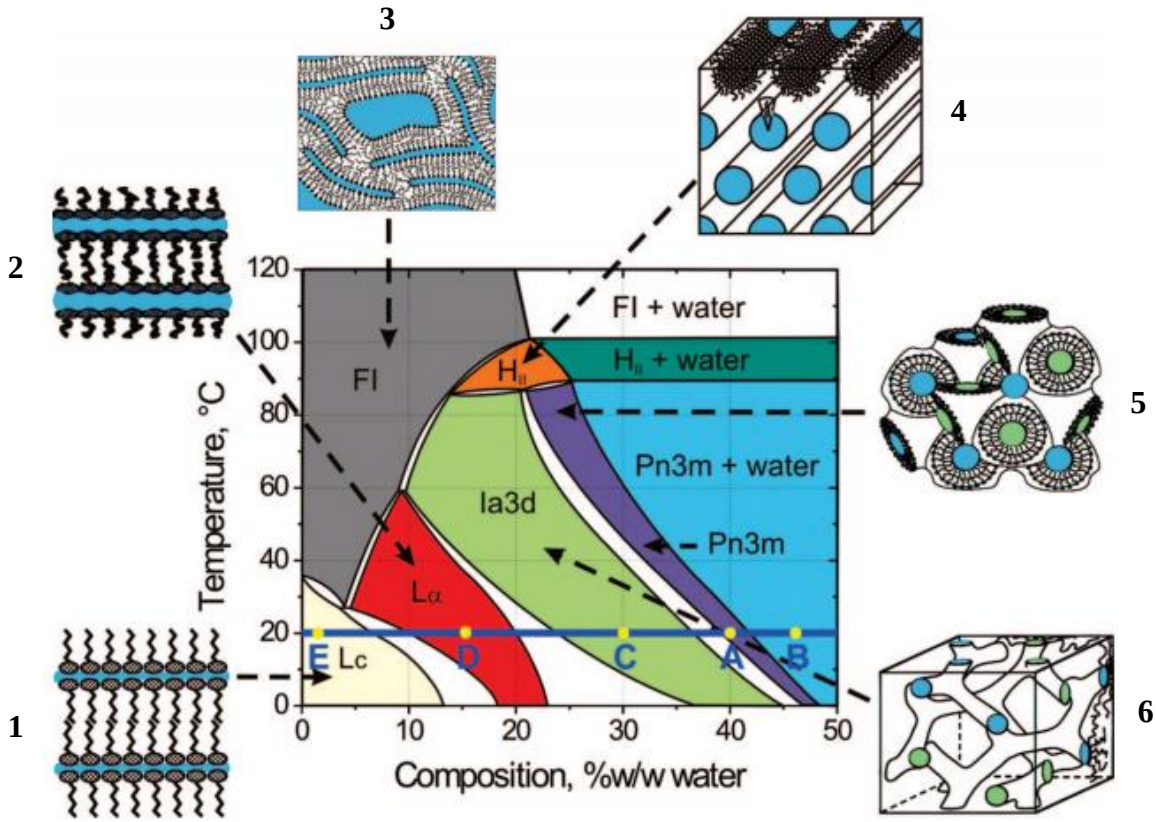


Figure 3.75. Monoolein/water phase diagram - temperature versus composition. The phase states and their corresponding cartoon representations. **1**. Lamellar crystalline. **2**. Lamellar liquid crystalline. **3**. Fluid isotropic. **4**. Inverted hexagonal. **5**. Cubic Pn3m. **6**. Cubic Ia3d. Diagram taken from Caffrey 2008.

Monoolein was selected as the supporting lipid and purchased from Sigma Aldrich and prepared using published methods (Caffrey and Cherezov, 2009). The protein of BAPKO_0422 and BB_0562 samples were reconstituted in Monoolein at 40% w/v using two syringes and a syringe coupler. Using a syringe dispensing unit 0.5µl of the lipid/protein mixture was ejected into each well of a 96 well LCP screening plate (Swissci) and overlaid with 0.5µl of precipitant solution before being sealed with UPV file for storage at 20ºC (section 2.2.9.2). In order to prevent evaporation and drying of the lipid/protein mixture plates were prepared and sealed in quarters.

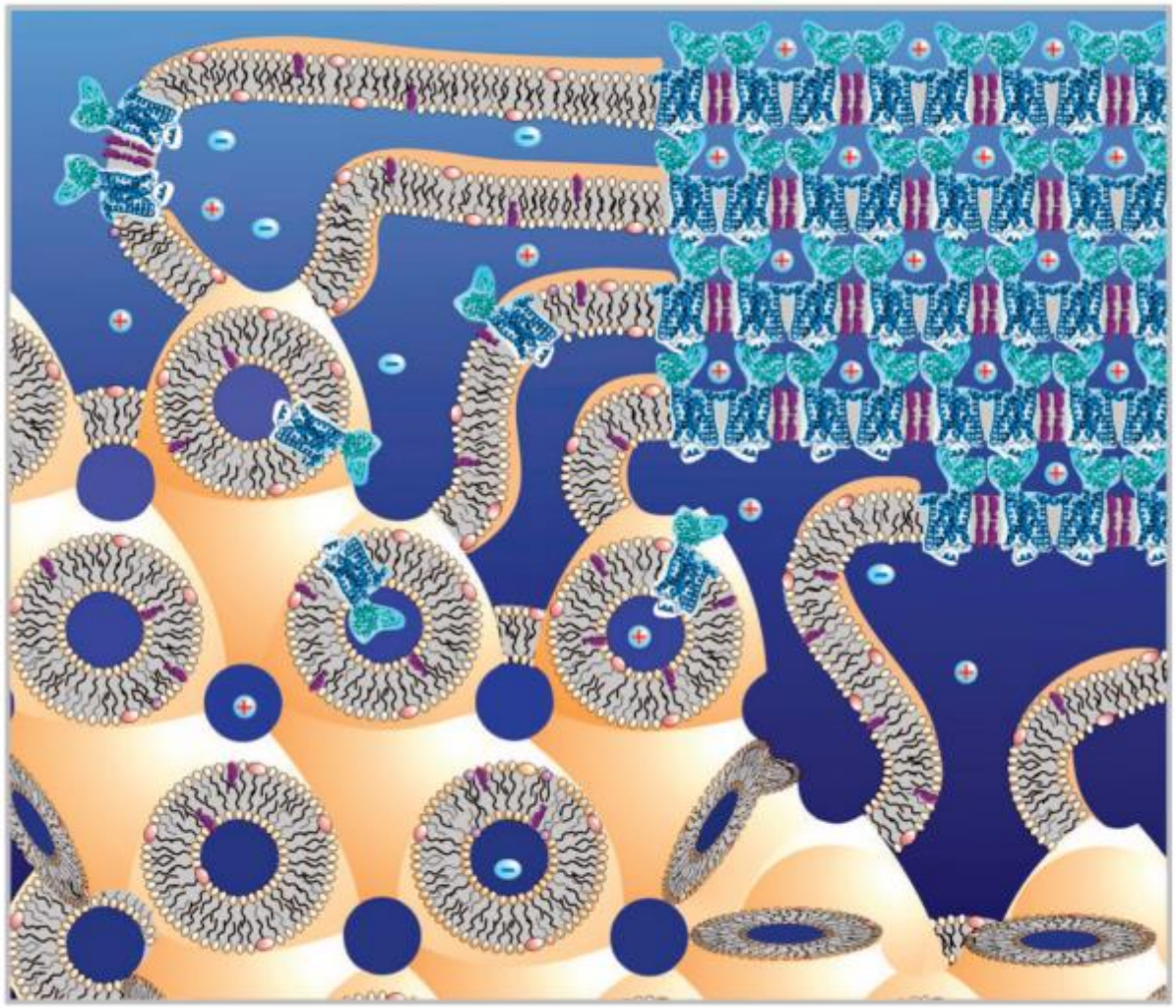


Figure 3.76. Formation of protein crystals within a lipidic cubic mesophase. A cartoon representation of crystal formation within the Pn3m cubic phase. Following reconstitution of the protein into the lipid bilayer environment as shown in the bottom left portion of the diagram the addition of precipitants causes a shift in the equilibrium of the cubic phase. This disruption leads to a phase separation where by protein molecules move towards a sheet-like or lamellar structure as shown in the upper right corner. Once several protein molecules congregate they begin to form regular arrays which may include detergent and lipid molecules. This organised crystal continues to grow from the continuous supply of protein molecules from the remaining cubic phase (Cherezov et al., 2006). Lipids are shown as light yellow ovals with tails, detergent molecules are shown as pink ovals with tail, protein molecules are shown in blue (pdb 2RH1 – Cherezov et al., 2007) and remaining native lipids are shown in purple. Figure taken from Caffrey 2008.

Alike with vapour diffusion trials plates were checked and photographed everyday for the first week and weekly thereafter. Under some conditions BAPKO_0422 appeared to be forming four forms of crystals as shown in figure 3.77. These crystals appeared small 3-dimensional facets (figure 3.77A), 2-dimensional triangular shaped bodies (figure 3.77B), 2-dimensional plate crystals (figure 3.77C) and 3-dimensional rods (figure 3.77D).

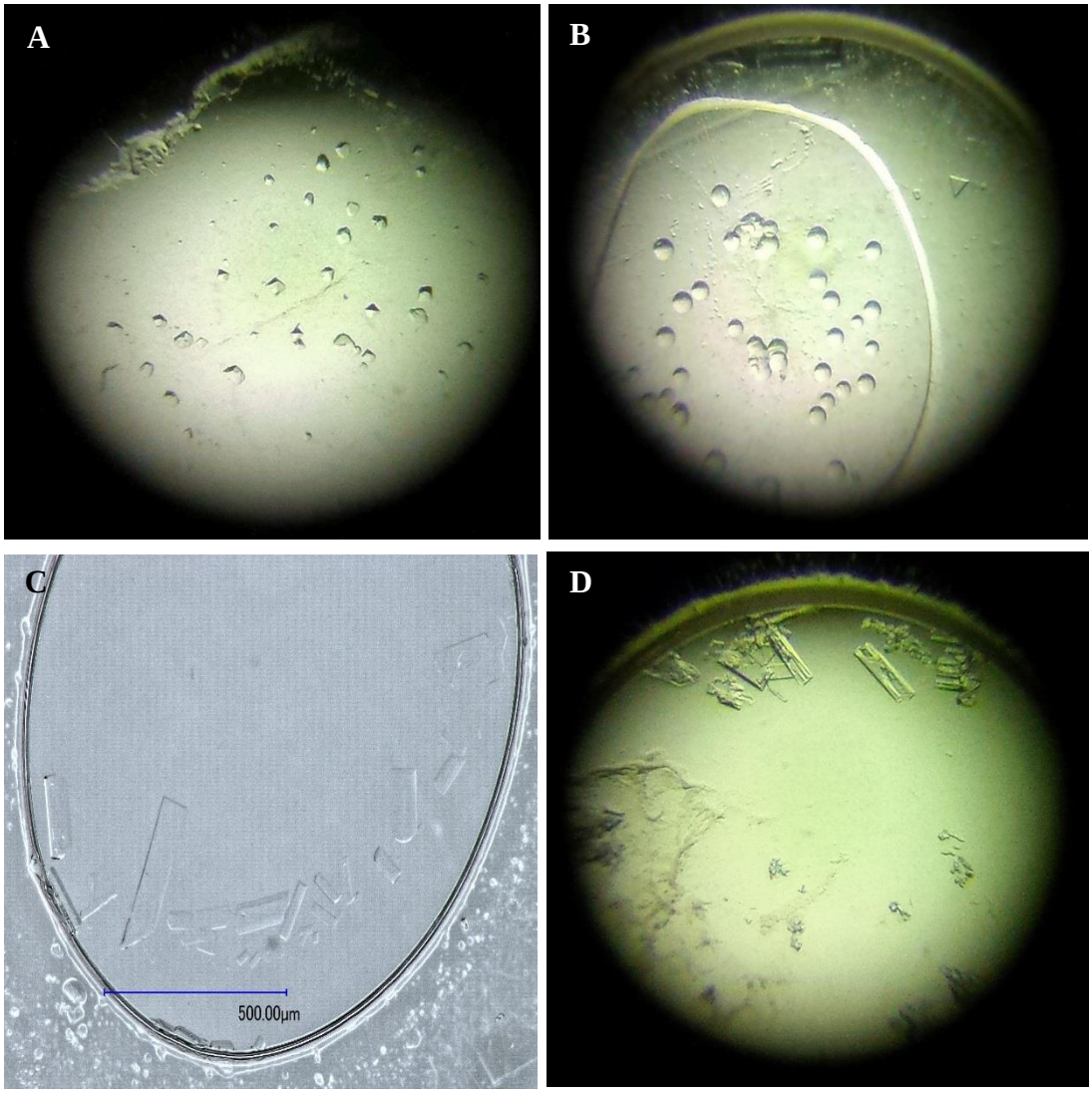


Figure 3.77. Lipid cubic phase crystallisation trails of *Borrelia* BAPKO_0422. Photographs taken after 7 days of incubation at 20°C. A. 3-dimensional facets which developed in 0.5 µl of lipid/protein mixture overlaid with 0.5µl of 10% v/v ethanol and 100mM CHES at pH 9.5. B. 2-dimensional triangular shaped bodies using 160mM ammonium sulphate, 47% v/v PEG 400 at pH 5.5 as the precipitant. C. 2-dimensional plate crystals using 47% v/v PEG 400, 170mM ammonium sulphate as the precipitant (photograph courtesy of Dyer 2013). D. 3-dimension rod crystals using 1M ammonium sulphate, 0.1M tris, pH8.5 as the precipitant.

Further investigation of the BAPKO_0422 crystals suggested that the 3-dimensional facets (figure 3.77A) and the 2-dimensional triangular bodies (figure 3.77B) could have been formed by the lipid monoolein due to a shift in the phase. As previously discussed the phase behaviour of the lipid is essential in order to promote crystallisation growth. This phase structure can easily be shifted to other phase orders by changes in the environment. 3-dimensional facets (figure 3.77A) are often associated with the Im3m cubic phase which can be induced by the presence of ethanol. In contrast the small triangular bodies (figure 3.77B) are often an indicator of the Pn3m phase. The instability of the lipid phase may have been detrimental to the LCP crystal trials and further work was carried out to ensure the set up and storage environment was constant and precipitants used were compatible. These facets and 2-dimensional triangular bodies were no longer observed suggesting that the phase had been stabilised and the changes were likely induced by components of the initial precipitants used. Several precipitant conditions during LCP crystallography of BAPKO_0422 gave rise to 2-dimensional rods and plates as shown in figure 3.77C. These rods were particularly difficult to manipulate, did not make a distinctive noise during crushing and appeared to dissolve when Izit dye was applied. These crystal were only observed in the presence of PEG and most often associated with 47% v/v PEG 400. It was not thought that these crystals were the result in a lipid phase shift nor did they share any characteristics with salt crystals. Attempts to remove them from the lipid phase with loops was problematic and their 2-dimensional structure was not suitable for diffraction. Attempts were made to optimise these conditions in order to produce smaller 3-dimensional crystals. However, no progress was made under the conditions tested.

Other crystals noted within the trials were large 3-dimensional rods as shown in figure 3.77D these were particularly suspicious as they were seen in most drops containing ammonium sulphate and were confirmed to be salt using Izit dye and crushing.

A number of limitations can be associated with the crystallisation experiments for BAPKO_0422 and BB_0562 which may need to be addressed in order to produce crystals suitable for diffraction experiments. Firstly, LDAO was used exclusively throughout the crystallisation experiments and may be responsible for the lack of success in obtaining suitable protein crystals. The use of different detergents can give rise to different crystal forms (Allen & Feher, 1984) and crystals of the highest quality may only be obtained from one or a small group of the numerous detergents available. An example of this was seen for bovine mitochondrial cytochrome c oxidase whereby diffraction of the crystals reduced when the chain length of the detergent was either decreased or increased beyond a

10-carbon acyl chain (Tsukihara et al., 1996). Future work should aim to screen a number of other stabilising detergents as research suggests this can be a major issue to overcome during crystallisation.

All crystallisation experiments were carried out using polyhistidine tagged BAPKO_0422 and BB_0562, this tag along with the amino acid sequence for the HRV3C cleavage site may have led to unwanted flexibility. Attempts were made to remove the tag using the HRV3C enzyme. However, digests were poor following numerous rounds of optimisation and tag removal was problematic. It is likely that the HRV3C cleavage site was no longer fully exposed once the protein was refolded. Future work should aim to produce further constructs to produce either untagged recombinant protein and purify by ion exchange chromatography or constructs which would express tagged protein with a larger span of amino acids between the HRV3C cleavage site and the start of the protein.

3.12. Discussion: *B. burgdorferi s.l.* OmpA-like proteins

The aims of the research project were to identify and characterise integral β-barrel proteins from the spirochaete *Borrelia*. Previous bioinformatic analysis sought to identify possible OmpA-like proteins from *B. burgdorferi s.l.* This initial work used hidden Markov models, fold prediction and signal sequence analysis in order to detect possible OmpA-like homologs which share very low sequence similarity. A total number of 12 proteins across *B. afzelii*, *B. burgdorferi* and *B. garinii* were identified to contain a possible OmpA-like domain with 9 of these containing a traditional signal sequence (Dyer, 2013). To date only around ten OMPs have been identified in *B. burgdorferi* (Brooks et., 2006, Lenhart & Atkins., 2010, Antonara et al., 2007, Bunikis et al., 2008, Coburn & Cugini et al., 2003, Cugini et al., 2003, Noppa et al., 2001, Skare et al., 1997, Wood et al., 2013 and Russell & Johnson., 2013) and none of these proteins fall into the OmpA-like family.

As the OmpA-like membrane spanning domain is well conserved amongst gram-negative bacteria it seemed reasonable to presume a homolog would be present within the *Borrelia sp*. Although spirochaetes differ somewhat from many other bacteria of the Gram-negative family they do share a number of common characteristics including the double membrane structure, presence of peptidoglycan within the periplasmic space and both groups express large numbers of lipoproteins. The more recent discovery of β-barrel assembly apparatus (Lenhart and Akins, 2010) and also secretion mechanisms within *Borrelia* highlight that the spirochaete has the ability to express, assemble, orientate and export membrane proteins into the outer membrane. Of *Borrelia's* known beta barrel membrane proteins none can be considered as similar to the well conserved OmpA. The presence of other beta barrel proteins within the spirochaetes outer membrane and *Borrelia's* ability to assemble such structures allows the assumption to be made that *Borrelia* may indeed express small beta barrel proteins such as the common 8-stranded beta barrel.

Scanning electron microscopy of *B. burgdorferi* has identified a number of integral transmembrane proteins and although the contribution of these to the membrane was much lower than seen in other gram negative bacteria such as *E. coli* it is safe to assume that there are more than those examples discussed in section 1.4.4 especially as a number of *Borrelia* genes have a recognisable signal sequence.

Much work has revolved around *B. burgdorferi*'s large number of lipoproteins many which confer favourable advantages with roles as key virulence factors (Cassatt et al., 1998, Haake. 2000, Schroder et al. 2003). Although these are important and often crucial for the spirochaetes long term survival within the host they are often non-essential for bacterial function and often encoded on *Borrelia's*

numerous plasmids. In contrast, all of the proposed integral membrane proteins within this study are encoded on the chromosome highlighting that they may play essential physiological roles.

The initial bioinformatics analysis undertaken by Dyer et al, Appendix 4, identified a number of potential membrane spanning beta barrel proteins across *B. burgdorferi*, *B. afzelii* and *B. garinii*, these proteins were all predicted as 8-stranded beta barrel proteins with signal sequences further supporting that this family were indeed integral membrane proteins with possible homology to the OmpA family. The subsequent aims for this thesis was to characterise a selection of these proposed OmpA-like proteins which are currently annotated either 'hypothetical' and 'conserved hypothetical' proteins.

Both BAPKO_0422 and BB_0562 were demonstrated to exist as a monomer in 0.3M NaCl, 50mM Tris, 0.1% v/v LDAO at pH8 this was confirmed by native PAGE (figure 3.13), SEC-MALLS (data not shown – Dyer, 2013), SEC (figure 3.14) and SAXS (section 3.10). In contrast BB_0406 and BG0408 were shown to form higher order oligomers including dimers, tetramers and octamers (figure 3.13).

The differences in behaviour observed between the four OmpA-like proteins indicate that although they are likely to share some level of homology they may have distinct roles. BB_0406 and BG0408 show the highest degree of sequence similarity of (90%). Both BB_0406 and BG_0408 contain two cysteine residues both of which are at position 31 and 40. Both proteins were seen to form oligomers. Species of monomeric, dimeric, tetrameric and octameric protein was found for both proteins during both reducing SDS-PAGE (figure 3.13) and native PAGE. These oligomeric species were confirmed by western blotting against the polyhistidine tag. However, the importance of these oligomers remains poorly understood and it is unknown whether these species are physiologically relevant. Literature searching for beta barrel proteins which are known to form octamers includes the staphylococcal γ-haemolysin toxin which is composed of the two different proteins LukF and Hlg2 and the octameric penicillin binding protein which is a homo-octamer of the pab87 protein. However *B. burgdorferi* is not known to encode any toxins nor is there any sequence similarity shared between these proteins.

Alternatively it could be proposed that BB_0406 and BG0408 are part of the lipocalin family as they share a number of lipocalin characteristics. Lipocalins are small proteins with a length of around 200 residues and molecular masses averaging 20kDa which transport small hydrophobic molecules. These proteins carry an N-terminal signal sequence and although sequence similarity within this family is low (~20%) their tertiary structures are well conserved. Lipocalins can be described as 8-stranded antiparallel β-barrels with both N and C-terminal helices. The barrel is open at one side and the structure encloses a binding pocket. They have been identified to form oligomers including dimeric

structures as observed in an odorant binding protein (Tegoni et al., 1996) and also complex octamers of the crustacyanins (Keen et al., 1991). The sequence similarity between lipocalins is low. However, many are seen to share three motifs yet others are observed to share only one yet the tertiary structure is still preserved.

Both BB_0406 and BG0408 contain one lipocalin motif known as the structurally conserved region 2 (SCR2) this motif is TDY- - I and is observed at the correct positioning (within the end of strand six) for both proteins (figure 2.10). Although lipocalins are known to share three SCR's some are known to only share one of these motifs yet retain the preserved lipocalin structure. Furthermore both BB_0406 and BG0408 are predicted to form an 8-strand antiparallel β-barrel and on some occasions homology modelling has shown the presence of an N-terminal helix. As both these proteins contain two cysteine residues and *in vitro* oligomers were observed for both proteins it was identified that these pair of *Borrelia* proteins could be part of the lipocalin family rather than the originally searched for OmpA-like family. As discussed, both proteins shared many of the lipocalin characteristics. However, crucially lipocalins are extracellular soluble proteins a characteristic which was not seen in either BB_0406 or BG0408 which are particularly hydrophobic and insoluble without the presence of harsh denaturants or detergents above CMC. Due to the behaviour of the four *Borrelia* OMPs it is unlikely any of these proteins belong to the lipocalin family and submission of their amino acid sequences to the online prediction software LipocalinPred resulted in negative scoring and no lipocalin prediction being made. The molecular envelopes for all four also did not support the notion that these proteins could be lipocalins with the final surface structures more closely resembling an elongated barrel alike those seen in transmembrane beta barrels as opposed to the short barrel structure seen in lipocalins.

The sequence similarity between the twelve *Borrelia* OmpA-like proteins identified and *E. coli* OmpA is very low and not statistically significant (less than 13%). However, this is not unusual for integral β-barrel membrane proteins which are often difficult to identify due to containing short membrane spanning strands (<10 amino acids) of alternating hydrophobic residues separated by loop regions which are often highly variable (Schulz, 2000). Proteins within the OmpA-like domain family (Pfam - PF01389) are made up of eight membrane spanning antiparallel β-strands connected by three short periplasmic turns and four large extracellular loops (Schulz, 2000). Existing structural data for several OmpA-like proteins suggest that the membrane spanning region is highly conserved and the highly variable extracellular loops may be responsible for the functional diversity of this protein family.

Sequence alignments of OmpA-like proteins with known structures show a high level of conservation of alternating hydrophobic residues separated by small side chain residues often glycine within the region spanning the membrane. This distinct alternation between these amino acid types is also seen within the proposed *Borrelia* OMPs (figure 2.10). Notably, within structurally characterised OmpA-like proteins aromatic residues such as tyrosine and phenylalanine can be identified at either end of the transmembrane spanning region. These aromatic residues are highly conserved and probably involved in the formation of the aromatic girdle (Ulmschneider & Sansom, 2001).

Multiple positively charged amino acids can be identified within the N-terminal region of all four of the *Borrelia* OmpA-like proteins investigated within the project. BAPKO_0422, BB_0406 and BG_0408 all contain an alanine in the predicted -1 position and proposed to be processed by a type I signal peptidase. However, this isn't observed for BB_0562. This may be an indication that this protein is not an integral transmembrane protein. However, both HMM searching and topology predictions (PRED-TMBB) suggest otherwise. As most of the proposed *Borrelia* OMPs from the early bioinformatics studies appear to contain an N-terminal signal sequence it is likely that these proteins utilise the Sec dependent secretory pathway in order to translocate across the inner membrane. As previously discussed BB_0562 does not contain an alanine in the -1 position, the bioinformatics study also identified this missing residue in BA0026 and BG0027. However, the close orthologue BB0027 contains the signal sequence. These differences may be unique to *Borrelia* signal sequences and somewhat different from the SignalP reference sets. Alternatively these subtly different signal sequences may have arisen, as *Borrelia* are known to encode three different type-I signal peptidases as opposed to the single protein found within *E. coli* (Paetzel et al., 2000). The presence of three such proteins may allow for some deviation from the normally conserved signal. Both BB_0406 and BG0408 have two cysteine residues at positions 31 and 40 and would normally indicate potential sites for N-terminal lipidation. Inspection of the N-terminal, however, is a poor match to other spirochaetal lipobox sequences (Setubal et al., 2006) and as these residues are predicted to be within a periplasmic turn they are likely to be disulphide bonded.

The C-terminal region of most integral outer membrane β-barrels appears to follow a signature sequence which resembles Z-x-Z-x-Z-x-Y-xF>, where Z represents non-polar residues and x represents any amino acid. The terminal residue is predominantly a phenylalanine (Struyvé et al., 1991) which appears conserved and has been proposed to be essential for the correct assembly of bacterial outer membrane proteins (Struyvé et al., 1991). Recent research by Gessmann et al., 2014, however, has highlighted that this terminal phenylalanine although acts to enhance processing by the

BamA apparatus is not essential and can be substituted by aromatic amino acids and more rarely by other non-polar amino acids. Deviations from the typical C-terminal signature and its recognition by BamA have been demonstrated to be species specific. An example is seen within Proteobacterial OMPs which have a positively charged residue in the penultimate position. However, this is never observed in *E. coli* (Robert et al., 2006). This C-terminal sequence is apparent in the majority of the proposed *Borrelia* OMPs identified in the bioinformatics study (Dyer, 2013) and can be detected in all four of the experimental *Borrelia* OMPs (BAPKO_0422, BB_0562, BB_0406 and BG0408). The most prominent deviation from the Proteobacterial C-terminal motif is the absence of phenylalanine at the C-terminus (with the exception of BB_0562) within the three experimental *Borrelia* OMPs this residue is replaced by a small non-polar residue such as isoleucine (BAPKO_0422) or the polar residue asparagine (BB_0406 and BG0408). This replacement of a terminal phenylalanine was also seen in the other 8 orthologues previously identified (Dyer et al., Appendix 4). In some cases the terminal residue was replaced by leucine.

In order to further investigate this replacement of the terminal residue an analysis of the C-terminal of other known *Borrelia* OMPs was undertaken (Dyer et al., 2015). The pore forming protein DipA contains a C-terminal motif similar to BAPKO_0422 and terminates with K-Y and the large P66 porin terminates with a polar sequence of S-G-S. Borellia BamA itself, a transmembrane β-barrel, terminates R-Y and the integral membrane protein P13 ends S-Y. It is therefore proposed that integral β-barrel membrane proteins found within *Borrelia* contain a *Borrelia*-specific C-terminal signature sequence which is recognised by the BamA apparatus. This sequence is proposed to resemble Z-x-Z-x-Z-x-Z-[KR]-[ILNY] again where Z represents a small non polar residue and x can be any amino acid. There appears to be a conserved positively charged residue, however, it is not always the penultimate residue. Rarely is the terminal residue a phenylalanine (with the exception of BB_0562). An investigation of *Borrelia* C-terminal residues of potential β-barrel proteins found a range of possible terminal residues including isoleucine, valine, lysine, asparagine and tyrosine.

Topology prediction and homology modelling of the proposed *Borrelia* OMPs demonstrate a lack of a C-terminal domain and it is expected that the C-terminal portion forms part of the terminal β-barrel strand, this region is also expected to contain a conserved C-terminal signal motif (Tommassen, 2010). As previously described this motif in bacteria is a 10-residue sequence consisting of three alternating non-polar residues at positions -9, -7 and -5 from the C-terminus. The side chains of these residues are in direct contact with the aliphatic region of the membrane. Within this signature

sequence the -3 residue is often a tyrosine and is involved in the aromatic girdle between the polar and non-polar environments (Struyvé et al., 1991).

In order to gain an insight into other possible roles of the *Borrelia* OMPs a literature search for homologues was undertaken. To date only two of the proteins identified as possible β-barrels have been investigated. One homologue BB_0405 has been demonstrated as surface exposed, present in outer membrane vesicles (Yang et al., 2011) and expressed in conditions representing both the tick and mammalian environment (Brooks et al., 2006). The second BB_0407 has been identified as a human factor H binding protein (Bhide et al., 2009) which initiated the factor H binding investigation.

Traditional cloning methods were utilised throughout the project in order to produce constructs to express recombinant *B. burgdorferi* OmpA-like proteins. The DNA supplied by Dr Volker Fingerle (German National Reference Centre for *Borrelia*) for *B. afzelii*, *B. burgdorferi* and *B. garinii* was of a high level of purity and nested PCR was not required in order to successfully amplify the target genes (section 3.2). The pET47 vector was selected for use alongside an *E. coli* expression system which allowed the addition of an N-terminal polyhistidine tag and a HRV3C recognition sequence between the tag and the start of the proteins. Early attempts to produce constructs resulted in the transformation of *E. coli* cells with empty plasmids. A large amount of pET47 was passing through cloning experiments uncut by both restriction enzymes. In order to rectify this, less plasmid was used within restriction digests, the amount of restriction enzyme was increased and the digestion time was extended. However, this failed to improve transformation yields. This was rectified by adding an additional restriction digest after the ligation step. This involved subjecting all ligation reactions to the restriction enzyme SalI. The SalI site was not present in any of the target gene inserts and the site is lost following the full digestion of pET47 with BamHI and NotI, making the site a desirable target for the destruction of any 'empty' vectors. Following the extraction and purification of the construct from successful clones samples were sequenced by Source Bioscience (Rochdale) using T7 promoter and terminator primers both forward and in reverse. The resulting sequences were then translated and subjected to protein BLAST (Altschul et al., 1990).

Recombinant protein for all four samples (BAPKO_0422, BB_0562, BB_0406 and BG0408) was produced using an *E. coli* expression system and was expressed as insoluble inclusion bodies. For the case of BAPKO_0422 and BB_0562 this protein was solubilised and refolded both by rapid dilution and on-column folding methods and gave a fair yield of up to 40%. This ease of refolding observed for BAPKO_0422 and BB_0562 reflects what has been reported for *E. coli* OmpA which readily

refolds in detergent. However, this was not shared by BB_0406 and BG0408 which often encountered final yields of 1-10%. In order to improve these yields future experiments should aim to refold using a better redox system or alternatively return to the cloning stages and mutate the two cysteine residues in hope to improve stability. Gel filtration of refolded BAPKO_0422 and BB_0562 normally gave a single peak corresponding to approximately 20kDa (section 3.4.5). Fractions preceding the peak were normally observed as higher order aggregates but this rarely appeared as a distinctive peak indicating the majority of the aggregated proportion was removed by high speed centrifugation. BB_0406 and BG0408 often eluted during gel filtration with several peaks which were normally poorly defined. The samples were fairly polydisperse with a mixture of high order aggregates, octamers, tetramers, dimers and monomers (Figure 3.14). This was often observed but at lower levels when under reducing conditions using fresh SEC buffer containing 10mM DTT with similar results occurring during reducing SDS-PAGE possibly indicating a fairly strong interaction between each protein unit.

Gel shift experiments demonstrated that all four *Borrelia* proteins were not modified by heat (section 3.5). Several membrane spanning β-barrel proteins have been demonstrated to conserve their folded state when solubilised in the detergent SDS at room temperature (Fairman et al., 2011). However, this is lost when exposed to boiling and the protein denatures. Gel shift assays are therefore useful and a shift in the apparent size of the membrane protein is observed when the protein sample is boiled. This however, was not seen for any of the four *Borrelia* proteins. This behaviour is consistent with another *Borrelia* outer membrane protein, P66 which is proposed to be a 24 stranded porin also observed to remain unmodified by heat (Kenedy et al., 2014). It has been suggested that this characteristic could be a general feature of OMPs from the *Borrelia* genus. As all four *Borrelia* proteins remained unmodified by heat, phase partitioning experiments were carried out to demonstrate that the proteins were amphiphilic. As Triton X-114 has a relatively low cloud point the detergents properties could be exploited in order to separate amphiphilic proteins from their soluble counterparts at a reasonable temperature (~23ºC). All four proteins were found within the detergent rich phase supplying evidence that these proteins could indeed be integral membrane components.

Circular dichroism is a powerful technique used to rapidly identify secondary structure, monitor unfolding behaviour and can be used to assess the binding properties of proteins. The technique is non-destructive providing the sample is not heated and requires only small amounts of often valuable samples. Throughout the project, CD was an essential tool to ensure protein samples were folded and to assess the percentage of secondary structure using deconvolution algorithms and reference data sets. Although powerful, CD has its limitations. Not only is it a low resolution technique but it is

sensitive to the composition of buffers and carry-over impurities. Buffer components such as high salt, detergents and some denaturants alongside high protein concentrations can cause a rapid increase in the H[T] voltage above a set threshold which differs for each piece of equipment. Once this H[T] voltage threshold has been met any data collected beyond this point is considered to be unreliable. It was therefore important to avoid any non-essential buffer components, keep detergent levels to a minimal and test several protein concentrations. The CDSSTR algorithm (Manavalan and Johnson, 1987) was selected for deconvolution as membrane proteins are underrepresented in all reference sets and this method would allow this bias to be over come to some degree. The CDSSTR method (Manavalan and Johnson, 1987) is a modification of the original variable selection method known as VARSLC which performs all possible calculations using a fixed number of proteins from the given reference set. The CDSSTR method (Manavalan and Johnson, 1987) can recognise proteins which are dissimilar from the test proteins and excludes them from the reference set. It has been suggested that this method produces the most accurate results but often takes the longest due to the number of calculations. The reference sets SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011) were selected purely as they contained the largest number of proteins the major difference between the two is that SMP180 (Abdul-Gader et al., 2011) contains an addition number of proteins often useful for the analysis of membrane proteins. The contents of the datasets and their corresponding pdb codes can be viewed in the appendix 2.

The CD experiments (section 3.9.2) revealed a classic β-strand spectrum with a maxima at 196 nm and the minima reached at 216 nm. The spectra generated agreed well with CD data from other structurally characterised 8-stranded β-barrels and secondary structure analysis predicted approximately 42% β-sheet, 4% helix, 12% turns and 40% unordered again agreeing well with other known integral membrane proteins.

The *ab initio* modelling from SAXS experiments (section 3.10) for the four *Borrelia* proteins revealed a peanut shaped structure. The structural dimensions of the molecular envelopes differed slightly but all four had a diameter close to 20-25 Å and a total length of 45-50Å. These parameters agreed well with the length and diameter of the homology models and are similar to other structurally characterised 8-stranded β-barrels. All of the structures are consistent with a monomeric barrel structure. The molecular envelopes for BB_0406 and BG0408 which both contain two cysteine residues were obtained from their reduced state and further work is required in order to elucidate whether these proteins are involved in any other high order structures. Participation of BAPKO_0422 and BB_0562 in a multimeric porin type structure. However, is unlikely and all experiments to date

have suggested these proteins exist as monomeric constituents within the outer membrane. This independent monomeric existence would not be unusual as *B. burgdorferi* BB_0405 has been identified as a monomeric surface exposed protein which does not make any significant interactions with any other major outer membrane proteins (Yang et al., 2011). Although the SAXS experiments provided valuable information with regards to the OMPs size and surface shape there are some limitations, firstly the deduced molecular envelope and the calculated sizes are likely to reflect some contribution from the detergent LDAO. Although extensive steps were taken in order to ensure the detergent concentration and the dispersity matched the sample it would be naïve to believe this is fully accounted for and it is likely that the final envelopes represent the *Borrelia* proteins with some level of interaction with the detergent LDAO. Secondly, as with many experiments involving membrane proteins there is likely to be some level of aggregation within the samples. This was normally identified by a small upturn at low q when visualising the scattering data. Monodispersity is considered crucial for bioSAXS experiments and large levels of aggregation can cause abnormal scattering which is difficult to account for. During this study it was found that low levels of such aggregation did not change the final model or calculated size on the condition that the upturn at low q did not obscure the overall curve and the estimated *Rg* agreed with previous samples. As a final note although two controls were used throughout the SAXS study (lysozyme and ribonuclease A) it would have been beneficial to use a membrane protein with a known structure as a further control in order to assess the contribution of detergent interactions. Unfortunately such a sample was not available but is being considered for future experiments. Although there are some limitations to the SAXS data the experiments provided a valuable insight into the size and shape of the *Borrelia* proteins and the acquired data further supported that BAPKO_0422, BB_0562, BB_0406 and BG0408 could indeed be small 8-stranded beta barrel proteins.

Crystallisation attempts were problematic, BB_0406 and BG0408 were considered unsuitable crystallisation targets due to their difficulty to refold, poor overall yield and their multiple oligomeric states. BAPKO_0422 and BB_0562 were notably stable at room temperature once folded and did not appear particularly susceptible to degradation and were selected for crystallisation experiments (section 3.11). Traditional vapour diffusion crystallography failed to produce any promising environments across a large number of conditions. This, however, was limited to the number of detergents screened as only LDAO was used throughout the project. A further influencing factor may be the presence of the poly-histidine tag and HRV3C protease recognition site which was not removed prior to crystallisation. This area of flexibility may have inhibited crystallisation or reduced protein-protein contact. However, it is not unusual to crystallise proteins with such tags still remaining

(Carson et al., 2007) . Further work should aim to screen other stabilising detergents such as DDM (n-Dodecyl-β-D-maltoside), OG (n-Octyl-β-D-maltopyranoside) and C8E4 (Octyltetraoxyethylene) which has been identified as particularly useful for the crystallisation of β-barrel outer membrane proteins from Gram-negative bacteria (Prive, 2007).

The complement system is a crucial component of innate immunity and it's activation by the alternative pathway is the initial obstacle for invading microbes (Sarma and Ward, 2011). In order to regulate complement formation, the regulator, factor H must differentiate between host and non-host surfaces. Procurement of factor H by pathogenic microbes is well reported and provides a level of protection against complement attack (Lambris et al., 2008). The overall effect of the acquisition of factor H to the microbial surface downregulates opsonisation which would normally involve the marking of the microbe for ingestion by a phagocyte. Alongside the evasion of phagocyte elimination the recruitment of host factor H prevents the amplification of the complement cascade which would lead to the formation of membrane attack complexes. The prevention of these processes is beneficial to any invading microbe but the evasion of the membrane attack complex is of particular importance to Gram-negative bacteria and spirochaetes (Meri et al., 2013). Microbial surface proteins have been identified to interact with several domains of the large glycoprotein. Both *Neisseria* (Madico et al., 2006) and group A *streptococci* (Blackmore et al., 1998) interact with factor H within domains 6-7. These microbial proteins are also known to bind the factor H like protein-1 (FHL-1) which is an alternatively spliced transcript from the FH-gene which contains domains 1 to 7 (Madico et al., 2006, Blackmore et al., 1998).

Some microbes including *B. burgdorferi* sensu stricto appear to interact with factor H at multiple sites (Kraiczy et al., 2004, Hellwage et al., 2001). The spirochaete protein CRASP-1 is known to bind to factor H via domain 7 (Kraiczy et al., 2004) in contrast to OspE which has been demonstrated to interact with domains 19 and 20 (Hellwage et al., 2001). The ability of *B. burgdorferi s.s.* to interact with factor H at multiple locations using a number of outer surface proteins increases the protection the spirochaete can mount against complement activation.

The *B. afzelii* protein BAPKO_0422 was identified as a human factor H binding protein *in vitro* (section 3.7 and Dyer et al., 2014 – Appendix 4). Binding occurred with incubations of human factor H at concentrations much lower than what is seen within a human host (<3 fold lower). Although BB_0562, BB_0406 and BG_0408 were not observed to bind human factor H they may play roles in the acquisition of factor H during the infection of other mammalian hosts. Factor H binding has been

observed as species specific and is worthy of further investigation. Further work is required in order to assess the interaction between BAPKO_0422 and human factor H possibly using isothermal titration calorimetry in order to ascertain a binding affinity and a co-crystal structure would be highly desirable. Alternatively mutant studies would be useful and allow some estimation of which BAPKO_0422 residues are essential for the interaction.

Proteoglycans are found within the extracellular matrix in all connective tissues and on the surfaces of several cell types they are highly glycosylated proteins composed of a core proteins and one or more GAG chain. These GAGs are classified into four distinct groups including, heparin, heparan sulphate, dermatan sulphate and chondroitin-6-sulphate. It has been understood for over a decade that *Borrelia* strains which can bind heperan sulphate appear to invade endothelial cells and those that can bind to the alternative GAG type, dermatan sulphate can attach to glial cells (Leong et al., 1998). Any strain that can target both these GAG classes have been demonstrated to invade both cell types. It is not unusual for pathogenic bacteria to target any of the four GAG molecules in their strategy to invade cells with GAG binding identified across the field of pathogenic bacteria.

Preliminary *in vitro* data suggests that the *B. afzelii* protein BAPKO_0422 may have some level of affinity for heparin. This feature was somewhat noticed by chance during the early stages of recombinant expression. As previously discussed in section 3.4.2 difficulties were encountered during the purification of BAPKO_0422. This recombinant protein failed to bind to a His-tag column due to an excess amount of DNA present within the sample. This DNA presence was unusual and wasn't seen for any other recombinant *Borrelia* protein. BAPKO_0422's affinity for DNA led to question whether the protein could bind other highly negatively charged molecules.

Due to time restraints *in vitro* heparin binding was only assessed using a HiTrap Heparin HP column (GE Healthcare) by means of a competitive elution gradient using 1M NaCl. BAPKO_0422 was retained on the column suggesting there is some level of interaction between BAPKO_0422 and heparin (section 3.8). However, as many proteins are known to bind heparin this interaction remains poorly understood and requires further investigation. Should this interaction be of physiological importance it may further our understanding of the mechanisms *Borrelia* employs to invade and disseminate though mammalian tissues. Interestingly biovar-specific differences in GAG binding have been observed in Chlamydia (Moelleken and Hegemann, 2007) and may also exist across *Borrelia* strains alike what has been suggested for the spirochetes acquisition of factor H.

The interaction between BAPKO_0422 and DNA could suggest a role in oligonucleotide transport. However, it is unlikely that BAPKO_0422 would function as a channel to transport DNA or RNA. If the homology model reflects the proteins true structure then the interior of the barrel appears too small for such transport. This does not rule out involvement within oligonucleotide transport. BAPKO_0422 could act as a smaller component in a more complex system responsible for the transport of DNA or RNA across the membrane.

Integral membrane proteins are the gatekeepers of the impermeable lipid bilayer which encases all living cells. They have a wide variety of roles including cellular transport (Verma et al., 2009, Oomen et al., 2004, Van den Berg et al., 2004, Chimento et al., 2003, Ferguson et al., 1998, Buchanan et al., 1999, Ferguson et al., 2002.), cell signalling (Srivastava et al., 2014), cellular architecture (Pautsch & Schulz, 1998) and enzyme activity (Vandeputte-Rutten et al., 2003, Snijder et al., 1999). The medical importance of this protein family cannot be overestimated with 50% of modern medicinal drugs targeting these surface exposed molecules (Adams et al., 2011). Although the determination of these protein structures has remained challenging it is crucial we push to further our understanding of these fascinating and medically important molecules.

In summary the data presented throughout the project suggests that *B. burgdorferi* sensu lato encode a number of proteins which could fall into the OmpA-like family (PF01389) with possible orthologues across several strains. As these proteins are encoded on the *Borrelia* chromosome they are likely to be play important roles and may have implications in virulence. The data suggest that BAPKO_0422, BB_0562, BB_0406 and BG0408 all form membrane spanning β-barrels and their topology prediction matches the domain defined by the Pfam family PF01389 with eight membrane spanning β-strands linked with short periplasmic turns and longer extracellular loops. This research supports the extension of the species distribution of the PF01389 domain to include bacteria from the Spirochete phylum (Dyer et al., Appendix 4). BAPKO_0422 is proposed as another of *Borrelia* numerous virulence factors and has been demonstrated to bind human factor H *in vitro* with factor H incubation levels much lower than what is observed *in vivo*. This acquisition of host factor H rescinds the complement response and possibly reduced antigenicity of BAPKO_0422's surface exposed loops.

Further work is required to investigate the host-specific factor H binding activity of the OmpA-like homologues and to determine other physiological functions of these proteins. The relationship between BAPKO_0422 and heparin requires further investigation *in vivo* in order to determine the

physiological importance of this interaction. On the whole furthering our knowledge of the outer membrane of *Borrelia* and its surface exposed epitopes may lead to the improvement of recombinant immunoblots currently used for diagnostic testing and the identification of new vaccine targets.

Chapter 4: Results & Discussion: *Borrelia* Superoxide dismutase A

4.1 Sequence analysis of *B. burgdorferi* SodA

Members of the spirochaete class can be divided into three distinctive families which include *Brachyspiraceae, Leptospiraceae* and *Spirochaetaceae* and within these families fall a number of disease causing members which include *B. burgdorferi*, the relapsing fever spirochaetes such as *B. recurrentis, Leptospira* which is responsible for leptospirosis or Weil's disease, *Treponema pallidum* the causative agent of treponematoses including syphilis and finally *Brachyspira pilosicoli* which can cause intestinal spirochaetosis. Prior to the initial sequence analysis it was assumed that most *Borrelia* proteins would share sequence similarity with other proteins expressed by bacteria from the *Spirochaetaceae* family. For example it was believed that *B. burgdorferi* SodA would be most similar to other SODs from *Leptospira* or *Treponema;* however, this was not the case.

As with many *Borrelia* genes the *bb0157* SodA gene is AT rich with a GC content of 24.6%. Nucleotide BLAST (Altschul et al., 1990) searches with exclusions on the *Borrelia* taxid returned no significant similarity which is not unusual for coding regions from the *Borrelia* genome. More sensitive protein BLAST (Altschul et al., 1990) searches, again excluding the *Borrelia* taxid identified a number of proteins with significant sequence similarity and are summarised in table 4.1 Intriguingly the proteins identified were mainly from the *Thermus* family, multiple sequence alignments of *B. burgdorferi* SodA and other similar proteins from *Thermus sp* are shown in figure 4.1.

When BLAST (Altschul et al., 1990) searching within the *Borrelia* taxid only, the *B. burgdorferi* SodA enzyme shared varying levels of sequence similarity to SODs in other *Borrelia* strains. The highest scoring SOD protein was from *B. finlandenis* with an identity of 99% and then down to a much lower identity of 67.3% when compared to *B. miyamotoi.* Considering how well conserved SODs generally are from humans to archaea this shift in sequence similarity from 99% to 67.3% between just two *Borrelia* strains highlights just how different these family members could be. Full BLAST (Altschul et al., 1990) search results are shown in table 4.1 and the corresponding multiple sequence alignment for *Borrelia* SODs can be seen in figure 4.2.

Description	Max Score	Total Score	Query Cover	E-value	Ident	Accession
Superoxide dismutase – *Thermus sp. 2.9*	221	221	97%	5e-69	51%	WP_039457669.1
Superoxide dismutase Mn – *Thermus thermophilus*	221	221	98%	6e-69	51%	WP_011172643.1
Superoxide dismutase – *Thermus aquaticus* (EC 1.15.1.1) Mn	221	221	98%	8e-69	51%	S07147
Superoxide dismutase – *Thermus oshimai*	221	221	98%	9e-69	51%	WP_016329036.1
Superoxide dismutase – *Thermus islandicus*	221	221	98%	1e-68	51%	WP_022798718.1
Superoxide dismutase – *Thermus filiformis*	220	220	97%	1e-68	51%	WP_038062973.1
Superoxide dismutase – *Chthonomonas calidriosea*	220	220	99%	2e-68	50%	CEK19605.1
Superoxide dismutase – *Thermus sp*. CCB US3 UF1	220	220	97%	2e-68	52%	WP_014515330.1
Superoxide dismutase – *Thermus thermophilus*	220	220	98%	2e-68	51%	WP_014629993.1
Superoxide dismutase - *Chthonomonas calidriosea*	220	220	99%	2e-68	50%	CEK19593.1

Table 4.1. Protein BLAST search results for *B. burgdorferi* SodA (BB_0157). The top ten hits following submission of the *B. burgdorferi* gene to protein BLAST (Altschul et al., 1990) with the *Borrelia* taxid excluded. BLAST (Altschul et al., 1990) searches were performed using the NCBI website (https://www.ncbi.nlm.nih.gov/) and all default settings.

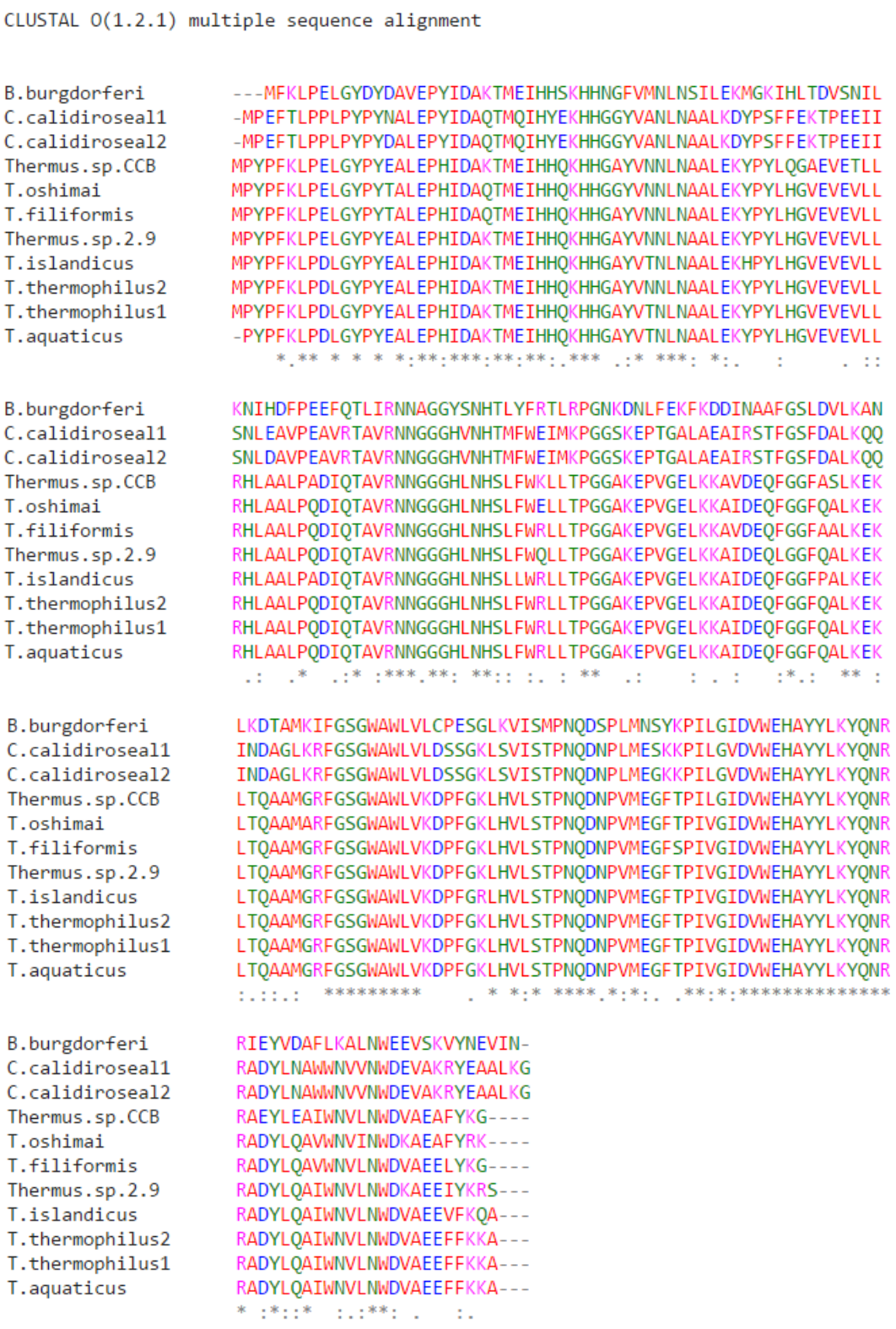


Figure 4.1. Multiple sequence alignment of *B. burgdorferi* SodA BLAST hits following the exclusion of the *Borrelia* taxid. The multiple sequence alignment was generated by Clustal Omega (Sievers et al., 2011) via the EMBL-EBI web service (http://www.ebi.ac.uk/services) using default settings and a text file of FASTA amino acid sequences supplied in the electronic supplementary data. All resulting BLAST (Altschul et al., 1990) hits were annotated as superoxide dismutases with accession codes shown in table 4.1. Pink residues are positive, blue are negative, red are non-polar and green any other.

Description	Max score	Total score	Query cover	E-value	Ident	Accession
Superoxide dismutase – *B. finlandensis*	416	416	100%	7e-149	99%	WP_008882325.1
Superoxide dismutase – *B. bissettii*	407	407	100%	2e-145	97%	WP_014023408.1
Superoxide dismutase – *B. garinii*	386	386	100%	4e-137	90%	WP_031478289.1
Superoxide dismutase – *B. spielmanii*	383	383	100%	6e-136	90%	WP_006433450.1
Superoxide dismutase – *B. afzelii*	383	383	100%	7e-136	90%	WP_044052137.1
Superoxide dismutase – *B. valaisiana*	381	381	100%	4e-135	89%	WP_006068873.1
Superoxide dismutase – *B. bavariensis*	380	380	100%	2e-134	89%	WP_011193502.1
Superoxide dismutase – *B. chilensis*	376	376	100%	5e-133	88%	AJA89990.1
Superoxide dismutase – *B. turicatae*	318	318	99%	5e-110	72%	WP_041178560.1
Superoxide dismutase – *B. hermsii*	317	317	99%	9e-110	74%	WP_025406281.1
Superoxide dismutase – *B. parkeri*	315	315	99%	3e-109	72%	WP_025375195.1
Superoxide dismutase – *B. coriaceae*	314	314	99%	1e-108	72%	WP_025407831.1
Superoxide dismutase – *B. crocidurae*	310	310	99%	3e-107	71%	WP_014696049.1
Superoxide dismutase – *B. anserina*	301	301	99%	1e-103	69%	WP_025419376.1
Superoxide dismutase – *B. recurrentis*	308	308	99%	3e-106	70%	WP_012538697.1
Superoxide dismutase – *B. duttonii*	308	308	99%	5e-106	70%	WP_012537922.1
Superoxide dismutase – *B. hispanica*	307	307	99%	6e-106	70%	WP_024655636.1
Superoxide dismutase – *B. persica*	303	303	99%	2e-104	69%	WP_024654104.1

Table 4.2. Protein BLAST search results for *B. burgdorferi SodA* (BB_0157) within the *Borrelia* taxid. The resulting sequence similarity following submission of the *B. burgdorferi* amino acid sequence to protein BLAST (Altschul et al., 1990) with the *Borrelia* taxid included. BLAST (Altschul et al., 1990) searches were performed using the NCBI website (https://www.ncbi.nlm.nih.gov/) and all default settings.

```
CLUSTAL O(1.2.1) multiple sequence alignment


B.bissettii         MFNLPELGYDYDAVEPYIDAKTMEIHHSKHHNGFVMNLNSILEKMDKIHLTDVSNILKNI
B.burgdorferi       MFKLPELGYDYDAVEPYIDAKTMEIHHSKHHNGFVMNLNSILEKMGKIHLTDVSNILKNI
B.finlandensis      MFKLPELGYDYDAVEPYIDAKTMEIHHSKHHNGFVMNLNSILEKMGKIHLTDVSNILKNI
B.chilensis         MFELPKLGYAYDAVEPYIDAKTMEIHHSKHHNGFVVNLNSIFEKMGKSHLKDVSTILKNI
B.spielmanii        MFKLPELGYNYDAVEPYIDAKTMEIHHSKHHNGFVINLNSILEKMGKSHLEDVSNILRNI
B.afzelii           MFKLPELGYNYDAVEPYIDAKTMEIHHSKHHNGFVINLNSILEKMGKSYLEDVSNILINI
B.valaisiana        MFKLPELGYNYDAVEPYIDAKTMEIHHSKHHNGFVMNLNSILEKMGKSHLKDVSDILKNI
B.garinii           MFKLPELGYNYDAVEPYIDAKTMEIHHSKHHNGFVMNLNSIFEKMGKSHLKDVSSILKNI
B.bavariensis       MFKLPELGYNYDAVEPYIDAKTMEIHHSKHHNGFVMNLNSIFEKMGKSYLKDVSSILKNI
B.anserina          MFKLPELGYAYDSLEPYIDAKTMEIHHSKHHNAYTVNLNSVLEKTEINYSQDIESILKNI
B.coriaceae         MFKLPELSYAYDALEPYIDAQTMEIHHSKHHNAYTVNLNSVLEKTELNCSKDIESILKNI
B.hermsii           MFKLPELGYAYDALEPYIDAKTMEIHHNKHHNAYTVNLNSVLEKTEINYSKDIESILKNI
B.turicatae         MFKLPELGYAYDVLEPYIDAKTMEIHHSKHHNAYTVNLNSVLEKTEINYSQDIESILKNI
B.parkeri           MFKLPELGYAYDVLEPYIDAKTMEIHHSKHHNAYTVNLNSVLEKTEINYSQDIESILKNI
B.persica           MFKLPELGYTYDALEPYIDAKTMEVHHTKHHNAYTVNLNSVLDQVKINYSHDIENILKNI
B.crocidurae        MFKLPELGYSYDALEPYIDAKTMEIHHSKHHNAYTVNLNSVLEKTKANCSYDIESILKNI
B.recurrentis       MFKLPELGYSYDALEPYIDAKTMEIHHSKHHNAYTVNLNSVLEKTKVNCSYDIESILKNI
B.duttonii          MFKLPELGYSYDALEPYIDAKTMEIHHSKHHNAYTVNLNSVLEKTKVNCSYDIESILKNI
B.hispanica         MFKLPELGYSYDALEPYIDAKTMEIHHSKHHNAYTVNLNSVLEKTKVNCSNDIESILKNV
                    **:**:*.* ** :******:***:**.****.:.:****::::       *:. ** *:


B.bissettii         HDFPEEFQTLIRNNAGGYSNHTLYFRTLRPGNKDNLFEKFKDDINAAFGSLDVLKANLKD
B.burgdorferi       HDFPEEFQTLIRNNAGGYSNHTLYFRTLRPGNKDNLFEKFKDDINAAFGSLDVLKANLKD
B.finlandensis      YDFPEEFQTLIRNNAGGYSNHTLYFRTLRPGNKDNLFGKFKDDINAAFGSLDVLKANLKD
B.chilensis         HNFPEEFQTPIRNNAGGYSNHTLYFRTLKPGNKSNLLEKFESDINAAFGSLEALKASLKD
B.spielmanii        QDFPDEFQTPIRNNAGGYSNHTLYFRTLKPGNKSNLLEKFENDINATFGSLDVLKAILKD
B.afzelii           HDFPDEFQTLIRNNAGGYSNHTLYFRTLKPGNKSNLLEKFENDINVTFGSLDVLKASLKD
B.valaisiana        HDFPNEFQTPIRNNAGGYSNHTLYFRTLKPGNKNNLLEKFANDVNTTFGGLDVLKASLKD
B.garinii           HDFPDEFQTPIRNNAGGYSNHTLYFRSLKPGNKSNLLEKFKNDVNTAFGSLDVLKASLKD
B.bavariensis       HDFPDEFQTPIRNNAGGYSNHTLYFRSLKPGNKSNLLEKFENDVNTAFGSPDVLKASLKD
B.anserina          QRFPKEFQPAIKNNAGGYSNHNLYFRTLKPGNRKNILGNFEEHVNATFGGLDNLKIVLKD
B.coriaceae         HRFPKEFQATIRNNAGGYSNHNLYFRTLKPGSKDNILESFEEHVNMTFGSLDNLKVALKE
B.hermsii           QRFPKEFQTAIRNNAGGYSNHILYFRLLKPGNKDNILENFEEHVNSTFGSLDNLKMVLKD
B.turicatae         QRFPKEFQAAIRNNAGGYSNHTLYFRILRPGNKDNILENFEEHVNATFGSLDNLKMALRD
B.parkeri           QRFPKEFQAAIRNNAGGYSNHTLYFRILRPGNKDNILENFEEHVNATFGSLDNLKMALRD
B.persica           QRFPAEFQKIVRNNAGGYVNHNLYFRTLKPGNKDNILKNFEEHVISTFGSLENLKVALKN
B.crocidurae        QRFPIEFQKVVKNNAGGYANHNLYFRTLKPGNKDNILKNFEEHVNATFGSLDNLKVALKD
B.recurrentis       QRFPIEFQKVVKNNAGGYANHNLYFRTLKPGNKDNILKNFEEHVNATFGSLDNLKVALKD
B.duttonii          QRFPIEFQKVVKNNAGGYANHNLYFRTLKPGNKDNILKNFEEHVNATFSSLDNLKVALKD
B.hispanica         QRFPIEFQKVVKNNAGGYANHNLYFRTLKPGNKDNILKNFEEHVNATFGNLDNLKMALKD
                      ** ***  ::****** ** **** *:**.:.*:: .* ..:  :*.  : **  *::
```

Figure 4.2. Multiple sequence alignment of *B. burgdorferi* SodA BLAST hits from within the *Borrelia* taxid. The multiple sequence alignment was generated by Clustal Omega (Sievers et al., 2011) via the EMBL-EBI web service (http://www.ebi.ac.uk/services) using default settings and a text file of FASTA amino acid sequences supplied in the electronic supplementary data. All of the BLAST (Altschul et al., 1990) hits were annotated as SODs with accession codes shown above in table 4.2. The dotted line separates Lyme disease causing strains (top) and relapsing fever strains (bottom). Pink residues are positive, blue are negative, red are non-polar and green any other.

The *B. burgdorferi SodA* gene was also analysed for the presence of rare codons, not commonly utilised within an *E. coli* expression system using the Rare Codon Calculator (RaCC), provided by the NIH MBI Laboratory for Structural Genomics and Proteomics UCLA (www.nihserver.mbi.ucla.edu/RACC/) with results shown in figure 4.3. As numerous rare codons were identified the gene was codon optimised and synthetically produced by MWG Operon prior to subcloning into the pET47 expression vector.

B. burgdorferi SodA (bb0157)

ATG TTT AAG CTG CCA GAA CTT GGT TAT GAT TAT GAT GCT GTT GAG CCT TAT ATT GAT GCT AAA ACT
ATG GAA ATT CAT CAT AGC AAG CAT CAT AAT GGT TTT GTA ATG AAT TTG AAT TCT ATT TTG GAA AAA
ATG GGG AAA ATT CAT TTA ACA GAT GTG TCA AAC ATA TTA AAA AAT ATT CAT GAT TTT CCA GAA GAA
TTT CAA ACT TTA ATA AGA AAT AAT GCT GGT GGT TAT TCT AAC CAT ACT TTA TAT TTT AGA ACT
TTA AGG CCA GGA AAC AAG GAC AAT CTT TTT GAA AAG TTT AAA GAT GAT ATT AAT GCA GCT TTT GGA
AGC CTA GAC GTT CTT AAG GCC AAT TTG AAA GAT ACT GCA ATG AAA ATT TTT GGA AGT GGT TGG GCA
TGG TTA GTA TTG TGT CCT GAG AGT GGC CTT AAA GTG ATT TCA ATG CCT AAT CAG GAT AGT CCT TTG
ATG AAT TCT TAT AAG CCG ATT TTA GGT ATT GAT GTT TGG GAG CAT GCC TAT TAC CTT AAA TAT CAA
AAT AGA AGA ATT GAA TAT GTT GAT GCA TTT TTA AAG GCT TTA AAT TGG GAA GAA GTT TCA AAA GTT
TAC AAT GAA GTG ATT AAT TAG

Figure 4.3. Rare codon analysis of the *B. burgdorferi SodA* gene. The *B. burgdorferi SodA* gene (*bb0157*) was analysed for rare codons using the Rare Codon Calculator (RaCC) provided by the NIH MBI Laboratory for Structural Genomics and Proteomics – UCLA (www.nihserver.mbi.ucla.edu/RACC/). A total of 8 rare codons were identified and are highlighted in red for rare Arg-codons, Green for rare Leu-codons and Blue for rare Ile-codons. No rare Pro-codons were found within the sequence.

4.2 Expression, solubility and the recovery of SodA from inclusion bodies

The *B.burgdorferi SodA (bb0157)* gene was codon optimised for expression within an *E. coli* system and synthetically produced (Eurofins MWG Operon). Solubility trials using several *E. coli* expression strains at varying conditions failed to yield soluble protein (section 2.2.7). However, high level expression as inclusion bodies was obtained and maximal expression levels reached after four hours of induction by IPTG at a final concentration of 1mM at 37^{o}C as shown in figure 4.5. Although autoinduction was tested for the expression of SodA, high levels of expression were obtained using the standard IPTG induced system and this method was carried forward.

In attempt to produce soluble SodA protein the vector was transformed into Rosetta DE3, NEB T7 and BL21 cells and expression tested using various concentrations of IPTG at temperatures ranging from 10 to 37^{o}C. No soluble protein was produced under the conditions tested with an example of the

SDS-PAGE analysis following expression at 10ºC with 1mM IPTG shown in figure 4.4. In order to assess the possibility of obtaining soluble protein the *B. burgdorferi* SodA amino acid sequence was evaluated by PROSO II an online protein solubility evaluator (Smialowski et al., 2006). PROSO II categorises proteins (without transmembrane regions) as either soluble or insoluble and scores the sequence between 0 (insoluble) to 1 (soluble). The method utilises over 80,000 known proteins in order to score the solubility of any given amino acid sequence. Analysis by PROSO II categorised *B. burgdorferi* SodA as insoluble with a score of 0.269 which may explain the difficulties encountered when attempting to recover soluble protein. Failure to obtain soluble *B. burgdorferi* SodA using an *E. coli* expression system led to the development of a refolding protocol as previously described in section 2.2.5. As refolding methods were already employed for the proposed *Borrelia* OMPs the same method was trialled for *Borrelia* SodA.

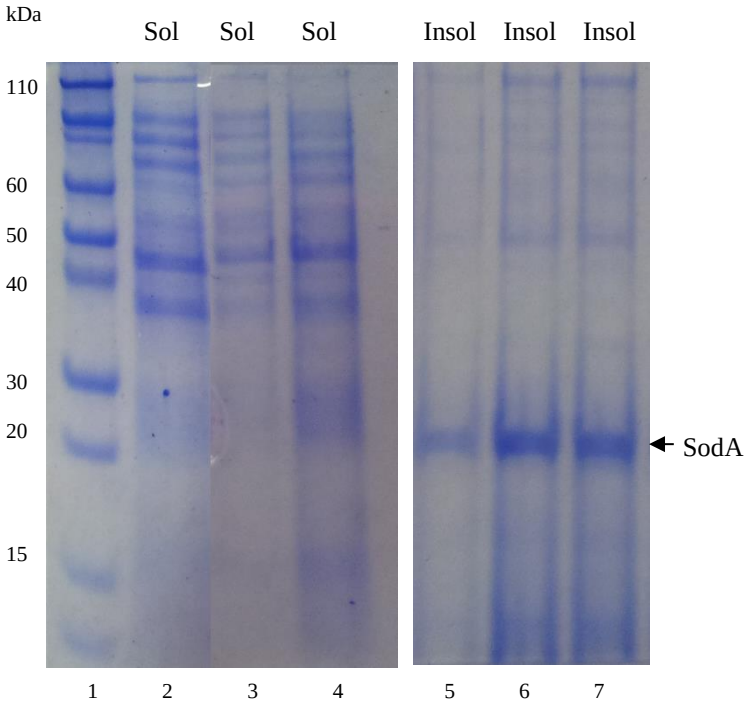


Figure 4.4. SDS-PAGE analysis of *B. burgdorferi* SodA expression. SDS-PAGE performed using 10% Bis/Tris gels and MES running buffer (50mM MES, 50mM Tris base, 0.1% w/v SDS, 1mM EDTA, pH 7.3) A total amount of 20µg of protein was mixed with LDS loading buffer (NuPage) and loaded per lane. BL21, NEB T7 and Rosetta *E. coli* cells were transformed with the SodA-pET47 construct and expression tested using 1mM IPTG at 10ºC. No soluble protein was produced and all of the recombinant protein was found within the insoluble pellet under the conditions tested. From left to right – Lane 1 – Novex Protein Standard, Lane 2 – Soluble fraction following expression of SodA using BL21 cells. Lane 3 - Soluble fraction following expression of SodA using NEB T7 cells. Lane 4 - Soluble fraction following expression of SodA using Rosetta cells. Lane 5 – Insoluble fraction following expression of SodA using BL21 cells. Lane 6 – Insoluble fraction following expression of SodA using NEB T7 cells. Lane 7 – Insoluble fraction following expression of SodA using Rosetta cells.

Following expression, cells were harvested by centrifugation, resuspended in lysis buffer (0.3M NaCl, 50mM tris base, 0.1% v/v triton X-100, 10mM DTT, 1mM EDTA at pH8) and the cell suspension separated into 15 ml aliquots. As previously described in section 2.2.4.1 lysis of the *E. coli* cells by pulsed sonication appeared complete following a sonication round of 6 minutes. This was tested by plating the lysate onto non selective LB agar and incubating overnight at 37ºC. Substantial *E. coli* growth was noted when the volume processed by sonication was greater than 15 ml indicating that the pulse was not propagating throughout the sample and therefore leading to incomplete lysis. Following cell lysis the soluble portion was removed following a round of centrifugation by removal of the supernatant which was retained for examination. Western blot analysis of the soluble lysate was negative for any his-tagged protein (data not shown) again indicating that *Borrelia* SodA forms insoluble exclusion bodies when using an *E. coli* expression system.

Development of the inclusion body wash method both improved the initial purity and increased the refolding yield, perhaps due to the removal of key contaminants which may promote aggregation during the refolding process. Following the removal of the soluble portion the remaining insoluble pellet was washed 6 times as described in the methods section 2.2.4.2. Unlike the *Borrelia* OMPs, recombinant SodA was not observed to interact with residual DNA and a DNase incubation prior to solubilisation was not required. Following the washing procedure the pellet was solubilised in denaturant (8M urea, 0.3M NaCl, 50mM tris base, pH8) overnight and the supernatant removed by centrifugation. By this point the solubilised recombinant protein was relatively pure considering it had not yet been subjected to any type of chromatography purification (figure 4.5B).

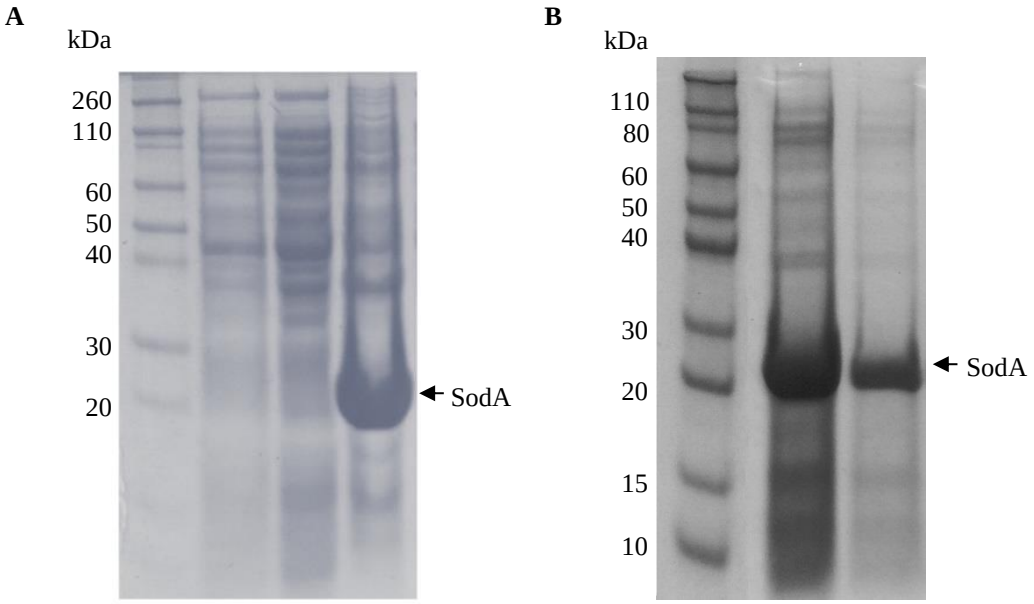


Figure 4.5. Expression and pellet washing of recombinant *B. burgdorferi* SodA from *E. coli* – SDS-PAGE performed using 10% Bis/Tris gels and MES running buffer (50mM MES, 50mM Tris base, 0.1% w/v SDS, 1mM EDTA, pH 7.3) **A.** SDS-PAGE analysis stained with coomassie blue. From left to right - Lane 1: Novex pre-stained protein standard, Lane 2: Pre-induction whole cell lysate, Lane 3: Soluble lysate, Lane 4: Insoluble fraction showing SodA monomeric protein. **B**. SDS-PAGE analysis of recombinant *B. burgdorferi* SodA pre and post wash. Lane 1: Novex pre-stained protein standard, Lane 2: recombinant SodA prior to pellet washing, Lane 3: recombinant SodA following pellet washing.

4.3 On-column refolding and purification of SodA

The washed recombinant SodA enzyme was efficiently purified and refolded by immobilised metal affinity chromatography (adapted from It et al., 2003). Initial attempts of refolding used a gradient from 8M urea down to no denaturant using an Akta Prime system. However, after further trials the yield improved when the gradient only ran as far 1M urea and the remaining denaturant removed during elution (section 2.2.7.4). These changes possibly aided refolding by stabilising hydrophobic interactions and suppressing aggregate formation during the final stages of refolding. Overall denatured SodA recovered from inclusion bodies and refolded on-column was successful with yields around 30-40% and an example chromatogram is shown in figure 4.6. Unlike the *Borrelia* OMP proteins, SodA faced no issues regarding to the binding of the nickel column under the condition that the selected buffers were above pH 7 so that the histidine residues were deprotonated and were able to bind to the nickel stationary phase.

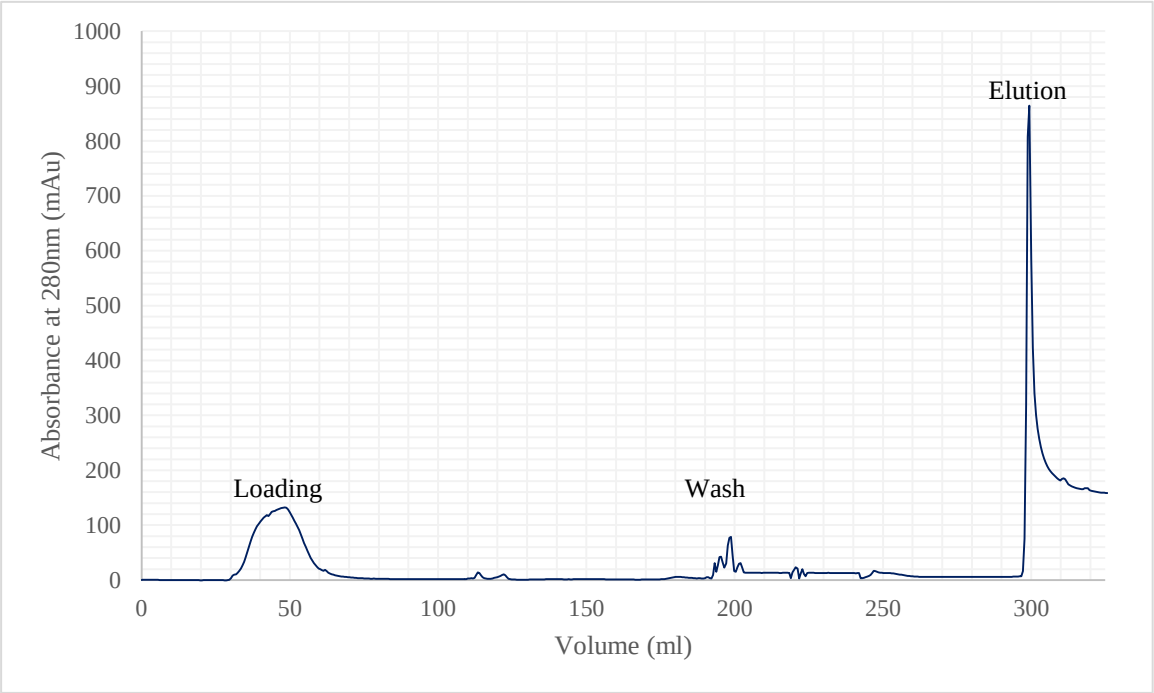


Figure 4.6. Immobilised metal affinity chromatography of *B. burgdorferi* SodA The blue line represents the changes in absorbance at 280 nm and recorded in mAu. Solubilised SodA was purified and refolded by IMAC using a Nickel column (GE Healthcare). The column was equilibrated as described in section 2.2.5.8 and solubilised SodA was loaded at a flow rate of 1 ml/min as represented by the peak in absorbance between 30 and 60 ml. Bound SodA was refolded by the gradient removal of the denaturant over a total volume of 120 ml at a flow rate of 0.5 ml/min prior to a wash step to remove non-specifically bound proteins (~200 ml). Refolded SodA was eluted by a gradient increase in the concentration of imidazole within the eluent (0.3M NaCl, 50mM tris, 0.3M imidazole, pH8) as shown by the increase in absorbance at 300 ml.

On-column refolding was achieved by the gradual removal of denaturant by means of a buffer exchange gradient. Protein yield was seen to increase when the refold gradient was both slower and a larger volume was used. However, this reached a maximum using a gradient length of 180 ml and a flow rate of 0.5 ml/min, totalling a refold time of 6 hours. Protein yield was also dependent upon temperature with the best results seen when refolding took place overnight and the room dropped to around 10ºC. Refrigerated refolding may have improved final protein yield further. However, this was not investigated (Xie & Wetlaufer, 1996). Following IMAC purification and on column refolding the recombinant SodA protein was eluted by a gradient increase of imidazole (0.3M NaCl, 50mM tris base, 0.3M imidazole, pH8) and fractions subjected to SDS-PAGE as shown in figure 4.7.

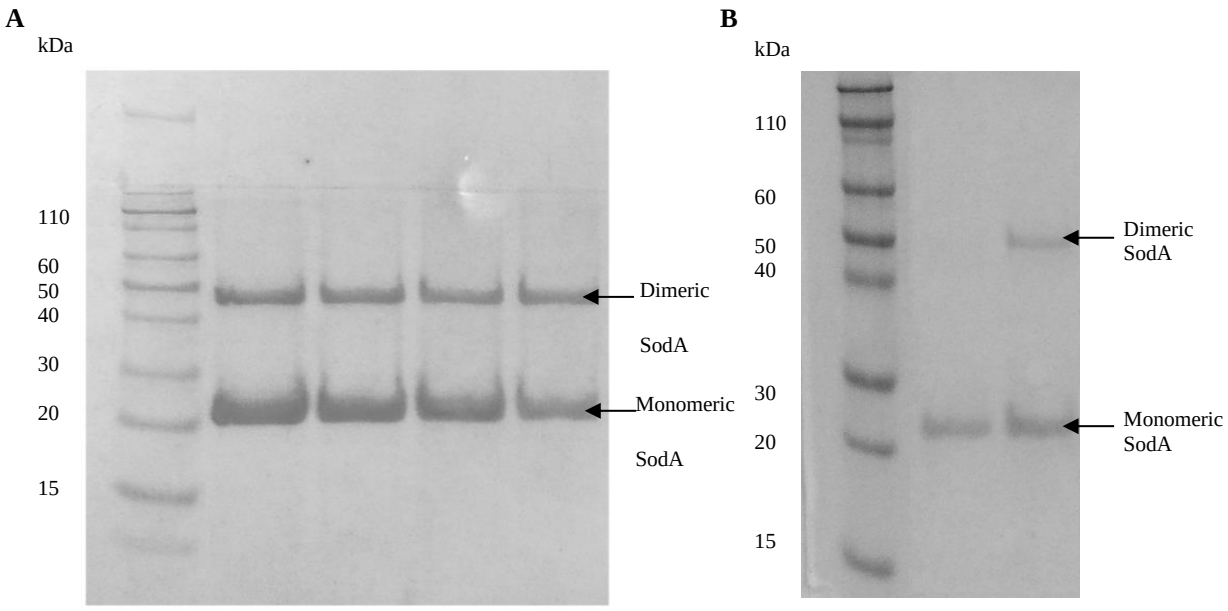


Figure 4.7. SDS-PAGE analysis of refolded and purified SodA. SDS-PAGE performed using 10% Bis/Tris gels and MES running buffer (50mM MES, 50mM Tris base, 0.1% w/v SDS, 1mM EDTA, pH 7.3) **A**. SDS-PAGE analysis of the SodA eluate from immobilised metal affinity chromatography. From left to right, 1. Novex pre-stained protein standard, 2-5 Eluted fractions following purification. **B**. Reduced and non-reduced SDS-PAGE analysis of purified SodA. From left to right, 1. Novex pre-stained protein standard, 2. Purified monomeric SodA from DTT reduction, 3. Purified monomeric and dimeric SodA.

Purity of the SodA extracted from inclusion bodies remained high from a single step purification method as shown by the SDS-PAGE analysis in figure 4.7. Fractions of purified SodA were pooled, concentrated and centrifuged to remove large aggregates before further polishing by SEC. SEC was conducted at a flow rate of 0.5 ml/min and 1 ml fractions collected. The molecular weight of each fraction was calculated using the calibration curve as shown in figure 2.5 in section 2.2.5.8. Monomeric SodA eluted within 12.5 ml giving a molecular mass of 23.5kDa less than the predicted molecular weight of 26.2kDa calculated using the online tool, ExPASy. Dimeric SodA as shown in the SEC chromatogram in figure 4.7 eluted within 11.2 ml. Using the calibration curve in figure 2.5 this volume represented a protein of 47kDa double the size of the apparent monomer. Fractions from both peaks were confirmed as recombinant *Borrelia* SodA by detection of the poly-histidine tag by western blotting and subsequent enzyme assays.

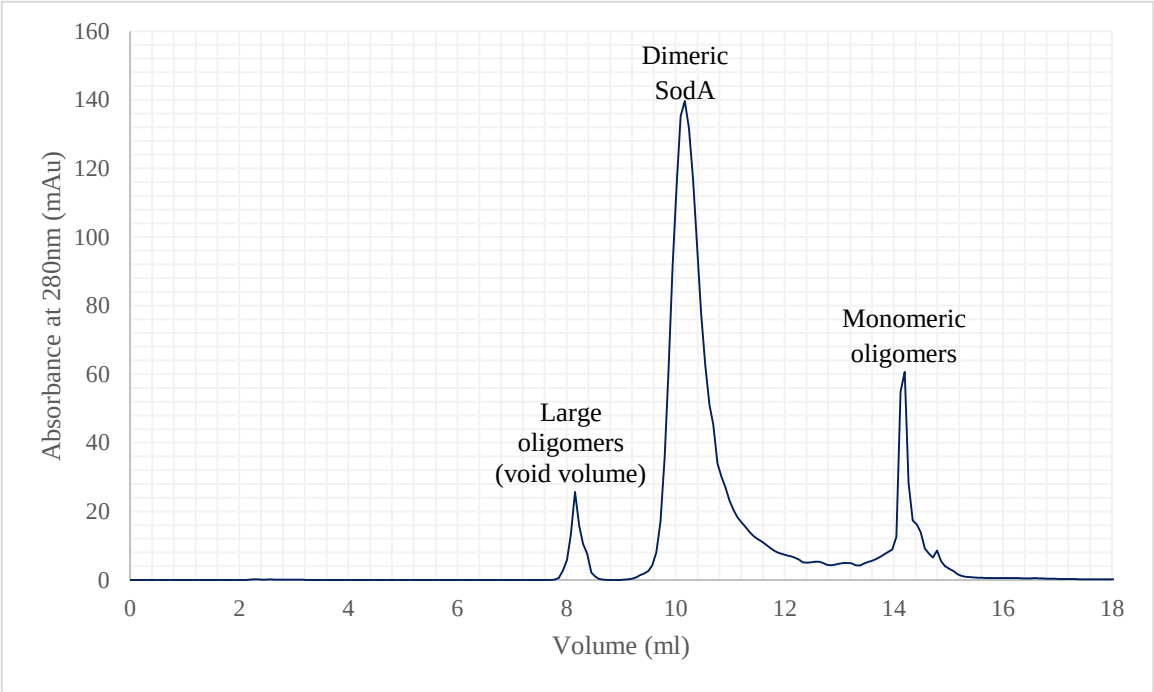


Figure 4.8. Final purification step of recombinant *B. burgdorferi* SodA by SEC. SEC was performed on a Superdex 75 10/300 GL column (GE Healthcare). The chromatogram shows three distinct peaks the first peaking at 8.1 ml following injection of the sample, this protein fraction fell within the void volume and is likely large aggregates. The second peak occurred at 11.2 ml and represents dimeric SodA at approximately 47kDa and the third peak at 12.5 ml corresponding with the 23.5kDa monomer. The exclusion volume/void volume was calculated as discussed in section 2.2.5.8 using blue dextran which eluted at 8.7 ml.

4.4 Preparation of Apo-SodA, Fe-SodA and Mn-SodA

Initial attempts to undertake SOD activity assays using the recombinant SodA protein produced as described in section 4.3, gave rise to multiple problems and adaptations to the existing methodology were required (section 2.2.7).

In order to determine enzymatic rates for *B. burgdorferi* SodA with bound Mn or Fe three samples were required, one containing no metal ions, the second containing only manganese and the third containing iron. Due to the severe implications of simply supplementing reactions with excess metal ions prior to the assay (further discussed in section 4.6) it became apparent that each sample would need to be produced independently. The first issue was to produce recombinant SodA with no metal co-factor (Apo-SodA). The simple use of chelators could not be used as these chemicals were found to interfere with the final SOD assay even in modest quantities and their effects could not be removed by the use of additional control reactions.

In order to produce Apo-SodA, denatured solubilised protein was stripped of its native metal ions by passing the sample through a stripped nickel column (GE Healthcare) twice. A third of this Apo-SodA was refolded as previously described with care taken when preparing buffers to prevent the addition of iron or manganese. For the case of Fe-SodA and Mn-SodA, Apo-SodA was supplemented with either 1mM $MnCl_2$ or $FeCl_2$ and refolded on-column as previously described.

The refolding of Apo-SodA was particularly problematic and gave a very low yield highlighting the possibility that the metal co-factor could be crucial for stability. Addition of $FeCl_2$ was also challenging due to its tendency to oxidise and following its addition the sample required filtration and rapid purification. All three samples were purified and refolded independently and the eluate pooled and centrifuged to remove large aggregates. The samples were not subjected to SEC due to the worry of sample loss, especially in the case of Apo-SodA. To ensure all three samples were folded prior to undertaking the SOD activity assay, samples were used for circular dichroism experiments further discussed in section 4.4.

4.5 Secondary structure determination of *B. burgdorferi* SodA

Circular dichroism experiments (section 2.2.8.1) were carried out to confirm whether recombinant *B. burgdorferi* SodA was folded. This was of particular importance for Apo-SodA which had proven difficult to refold. Secondly, the experiments aimed to estimate the amounts of secondary structure elements in order to compare to other known superoxide dismutases. Circular dichroism experiments were carried out on a JASCO J-810 spectropolarimeter (University of York) and the resulting data analysed using the online server dichroweb (Lobley et al., 2002).

Four samples were used for circular dichroism experiments. These included, *B. burgdorferi* SodA with endogenous metal ion (this sample had no alteration made to the existing metal ion and its co-factor was unknown and obtained during recombinant expression by *E. coli*), Apo-SodA, Mn-SodA and Fe-SodA with the resulting spectra shown in figure 4.9. Although only one concentration for each SodA sample is shown within figure 4.9, several concentrations were used for CD in order to provide the clearest spectra within a suitable mdeg range. Data for all concentrations can be seen within the appendix. CD data was recorded from 260 nm down to 185 nm and all data excluded once the H[T] reached 600V. All experimental data were submitted to the online server DichroWeb (Whitmore and Wallace, 2004, Whitmore and Wallace, 2008) using the CDSSTR algorithm (Compton & Johnson, 1986, Manavalan & Johnson, 1987) with reference sets SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011). The full deconvolution results are shown in table 4.3 with the experimental and predicted fit shown in figure 4.10. As several samples were used for CD and two independent reference sets used for deconvolution the final percentages of secondary structure were averaged. Mn-SOD was deemed to be composed of around 57% helix, 14% β-strand, 12% turns and 17% disordered, Fe-SOD was calculated as 60% helix, 16% β-strand, 13% turns and 21% disordered and Apo-SOD as 55% helix, 11% β-strand, 13% turns and 21% disordered. The deconvolution for all three samples, Mn, Fe and Apo-SOD agreed well and was in line with literature data for *T. filiformis* SOD with 58% helix (Mandelli et al., 2013) and *T. thermophilus* HB8 SOD (Sato & Nakazawa, 1978) with 56% helix, both of which had previously been identified to share significant sequence similarity.

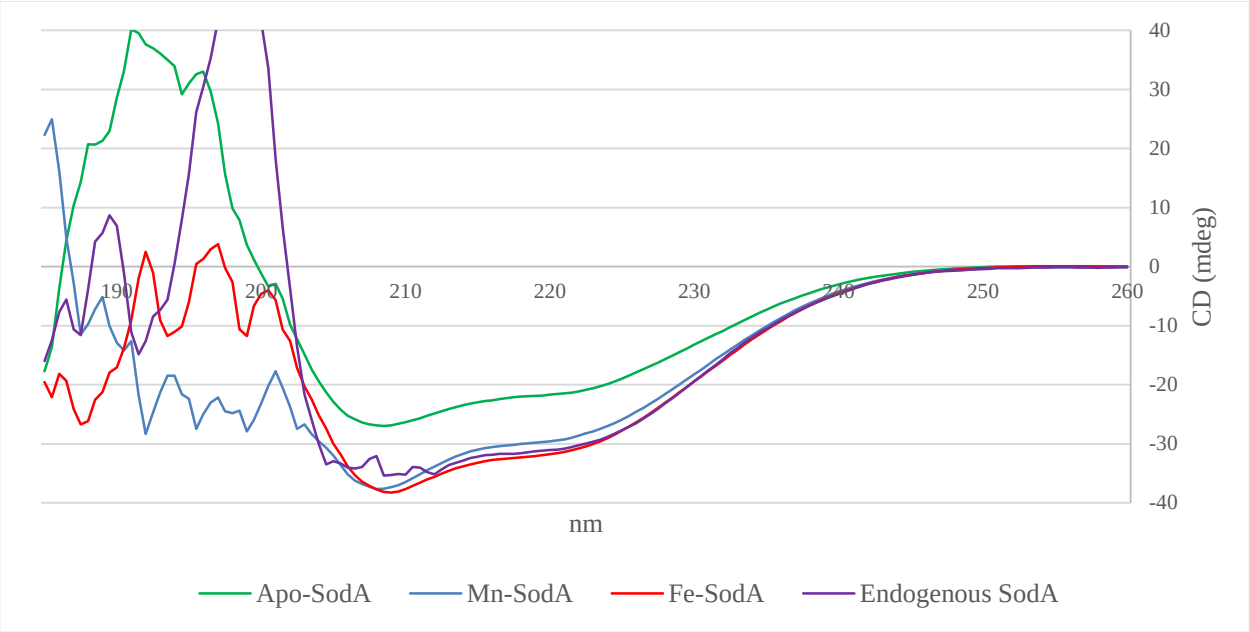


Figure 4.9. Circular dichroism of *B. burgdorferi* SodA at 20ºC. The recorded CD profile for four samples of *B. burgdorferi* SodA in 0.3M NaCl, 50mM tris base, pH8 at 0.01mg/ml. Data was acquired over 260-185 nm at 0.5 nm increments. For Mn-SodA, Fe-SodA and endogenous SodA the H[T] voltage crossed 600V at 204 nm and for Apo-SodA the threshold was reached at 199nm, data exceeding this points was discarded prior to deconvolution. The four profiles are all indicative of a folded protein with a large proportion of helix. Apo-SodA is seen to differ from the other three samples however the profile does follow the same pattern indicating a high amount of helical structure. The Apo-SodA sample was later found to be at a slightly lower concentration than the other samples at 0.008mg/ml explaining the deviation.

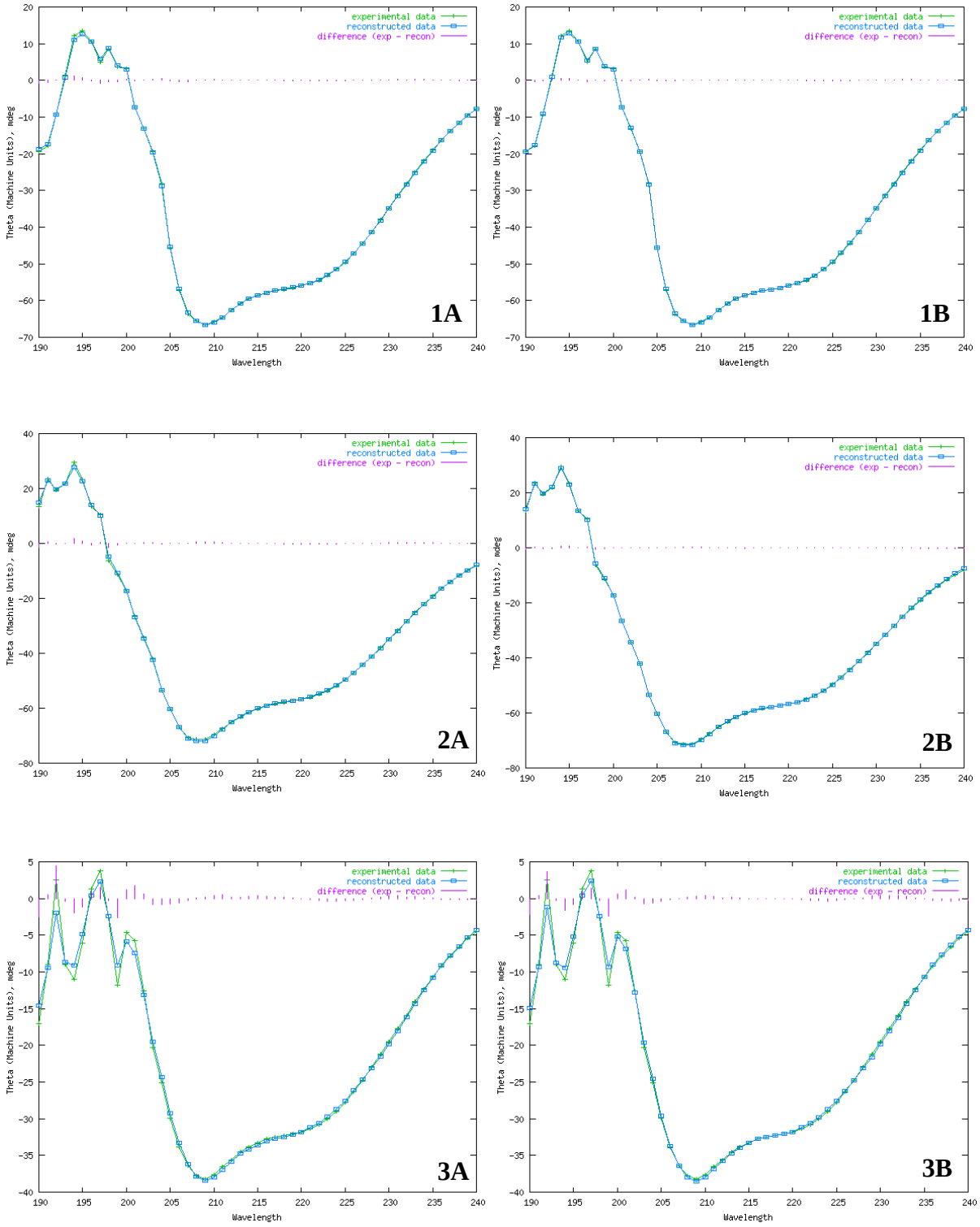


Figure 4.10. CDSSTR analysis of *B. burgdorferi* SodA samples. CD data for apo, Mn and Fe-SodA were analysed by dichroweb using the CDSSTR method (Compton & Johnson, 1986, Manavalan & Johnson, 1987) and data sets SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011). The green data is the experimental CD values, the blue is the reconstructed curve and the purple is the difference between the two. For all samples the difference between the experimental and the reconstructed values is minimal with the exception of the data at low nm where the HT voltage exceeded 600 in turn reducing the values accuracy. The left spectra were produced using the SP175 (Lees et al., 2006) data set and the right spectra were generated from the SMP180 (Abdul-Gader et al., 2011) data set. From top to bottom, 1A – Apo-SodA (SP175), 1B – Apo-SodA (SMP180), 2A – Mn-SodA (SP175), 2B – Mn-SodA (SMP180), 3A – Fe-SodA (SP175), 3B – Fe-SodA (SMP180).

Protein Sample	Method: CDSSTR - Reference set SP175					Method: CDSSTR - Reference set SMP180				
	Helix	β-strand	Turns	Unordered	NRMSD	Helix	β-strand	Turns	Unordered	NRMSD
Apo-SodA (0.2 mg/ml)	0.58	0.13	0.12	0.17	0.003	0.56	0.13	0.1	0.2	0.003
Apo-SodA (0.1 mg/ml)	0.58	0.10	0.15	0.17	0.036	0.50	0.08	0.14	0.28	0.030
Mn-SodA (0.2 mg/ml)	0.59	0.15	0.09	0.17	0.002	0.59	0.18	0.08	0.15	0.002
Mn-SodA (0.1 mg/ml)	0.55	0.13	0.11	0.2	0.003	0.55	0.13	0.13	0.19	0.002
Fe-SodA (0.2 mg/ml)	0.62	0.14	0.09	0.15	0.002	0.59	0.16	0.09	0.16	0.001
Fe-SodA (0.1 mg/ml)	0.63	0.13	0.07	0.18	0.002	0.59	0.19	0.08	0.14	0.002
Native SodA (0.2 mg/ml)	0.63	0.09	0.13	0.15	0.006	0.58	0.14	0.12	0.16	0.002
Native SodA (0.1 mg/ml)	0.62	0.013	0.07	0.18	0.002	0.59	0.2	0.08	0.13	0.002

Table 4.3. Deconvolution data for *B. burgdorferi* SodA. CD spectra were obtained for Apo-SodA, Mn-SodA, Fe-SodA and Native SodA across several concentrations. Two data sets for each sample were then processed by dichroweb for assessment of the secondary structure components. Deconvolution was undertaken using the CDSSTR method (Compton & Johnson, 1986, Manavalan & Johnson, 1987) using both SP175 (Lees et al., 2006) and SMP180 (Abdul-Gader et al., 2011) as independent data sets.

4.6 Small angle scattering experiments for *B. burgdorferi* SodA

Samples of recombinant *B. burgdorferi* were prepared as explained in section 2.2.8.2 and a dialysis matched buffer blank prepared in the same way. Scattering data were recorded both in-house on a Bruker NanoStar and at Diamond Light Source B21 (DLS funded – Session ID - SM12308). Data from in-house experiments showed a large upturn at low q indicative of sample aggregation although the protein sample had behaved well during SEC and had eluted at the expected volume. It was postulated that the sample had aggregated during the 2400s data collection period possibly due to the slight increase in temperature within the vacuum chamber. All scattering data was deemed unsuitable for further analysis due to the high level of aggregation.

Further SEC samples were transported on ice for data acquisition on beamline B21 at DLS. SodA samples were centrifuged for 30 mins at 16,000 x g at 4^{o}C to remove large aggregates and sample aggregation assessed at 340 nm. At this point the sample was considered suitable for data collection and loaded at various concentrations into the automated sample holder. The sample movement was monitored using the on-board capillary camera and at first exposure the sample precipitated out of solution. As the temperature was stabilised at 20^{o}C it was unlikely that temperature was responsible as previously thought and it is likely that the protein responded negatively to X-ray exposure possibly due to oxidation of the metal ion. This precipitation/aggregation was also seen with the addition of stabilisers such as 10% v/v glycerol and all further SAXS experiments were terminated.

4.7 SOD activity and chemical characterisation

B.burgdorferi SodA was originally characterized as a Fe-SOD based upon its resistance to cyanide yet enzymatic sensitivity to H_2O_2 (Whitehouse et al., 1997). Three years later a study demonstrated that this classification for *Borrelia* was unlikely. Their reasoning included that *Borrelia* have an extremely low intracellular iron concentration with evidence suggesting this could be as low as 10 atoms per cell (Posey and Gherardini., 2000). Furthermore, *Borrelia* have no growth requirement for iron and there is a distinct lack of cellular machinery for the uptake of iron (Posey and Gherardini., 2000). Further studies failed to confirm Whitehouse's findings that *B.burgdorferi's* SOD was enzymatically sensitive to H_2O_2 but resistant to cyanide however instead suggested that the *B.burgdorferi* SOD was resistant to both treatments and was indeed a Mn-SOD (Troxell et al., 2012).

SOD activity was assessed by the use of a SOD determination kit (Sigma 19160) which allows SOD activity to be determined based upon Dojindo's water-soluble tetrazolium salt, WST-1 (2-(4-Iodophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H tetrazolium, monosodium salt) which produces a water soluble formazan dye upon reduction with a superoxide anion. The rate of reduction with O_2 is linearly related to xanthine oxidase activity which is inhibited by SOD allowing inhibitory activity to be determined by use of a simple colorimetric assay alongside appropriate blank controls.

Initial activity testing (section 2.2.7.5) highlighted a number of problems which needed addressing. Originally the assay was performed by supplementing recombinant *B. burgdorferi* with either 1mM $MnCl_2$ or 1mM $FeCl_2$. However, excess amounts of these two additives terminated the colorimetric reaction. Also the rapid oxidisation of $FeCl_2$ gave rise to other colour changes within the reaction which couldn't be removed simply by a suitable blank or control.

Apo-SodA produced as described within the methods section gave an averaged 3% inhibition shown in figure 3B. This indicated that the method used to produce apo-protein was relatively successful and the majority of SodA protein was stripped of any native co-factors. Apo-protein supplemented with a supply of $FeCl_2$ prior to refolding also failed to inhibit the colorimetric assay to a reasonable level, giving an averaged inhibition of 5.9%. Apo-protein supplemented with $MnCl_2$ and refolded using the same method as Fe-SodA gave a significantly higher percentage inhibition at 82.9%.

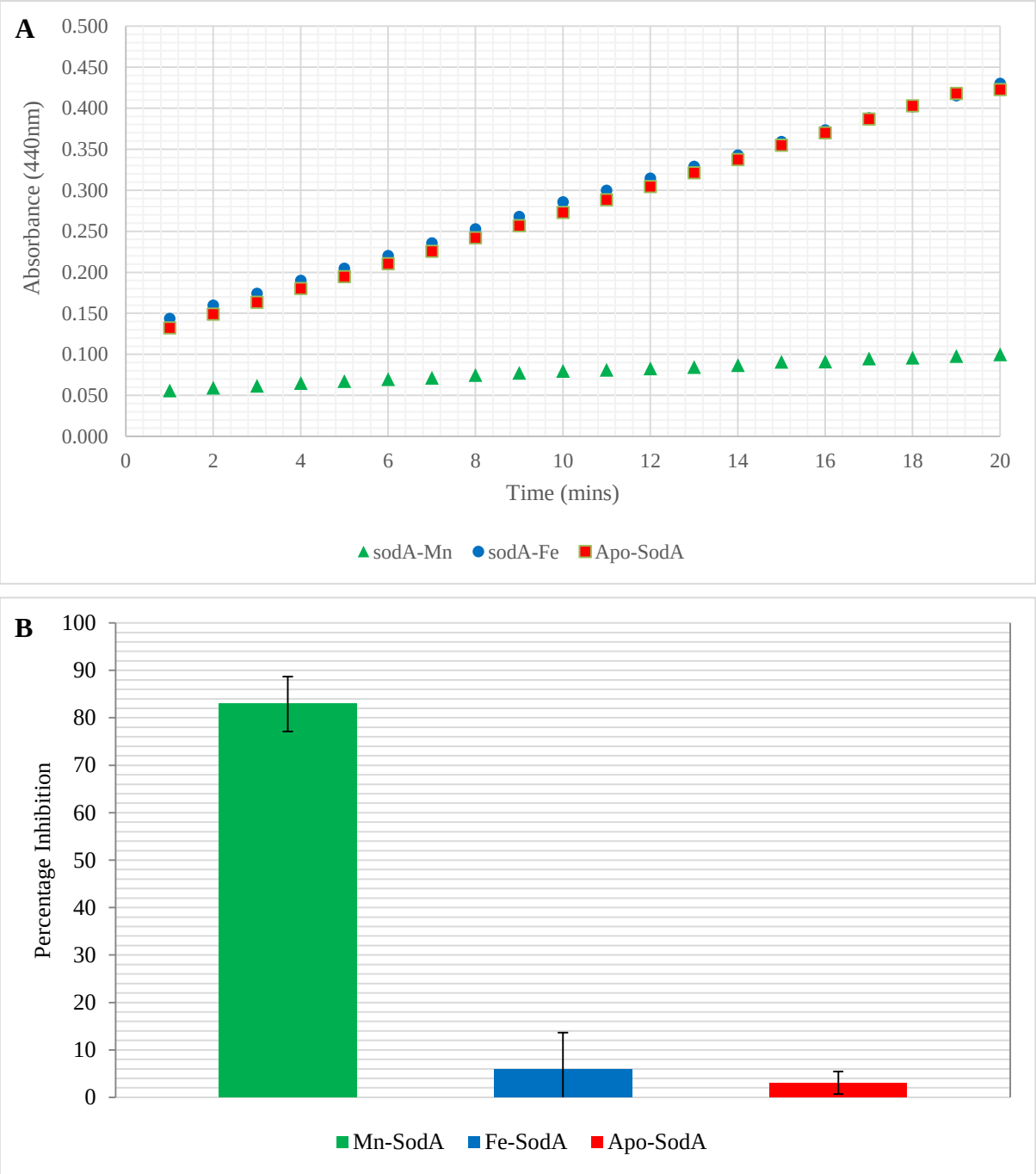


Figure 4.11. SOD activity and inhibition curves. Assays were performed under manufacturer's instructions and SodA samples were prepared in 0.3M NaCl, 50mM tris base, pH8. Error bars were calculated as standard error.

A. Assay inhibition curves for *B. burgdorferi* SodA samples. The absorbance was monitored at 440 nm and values recorded every minute for a total of 20 minutes. Data was averaged from 4 repeats and the background subtracted.

B. Percentage inhibition of the colorimetric SOD assay. Data is averaged over 4 repeats and the background subtracted. Percentage inhibition was calculated from an end point of 20 minutes. Bars from left to right, GREEN – Apo-SodA supplemented with a supply of $MnCl_2$ prior to refolding, total percentage inhibition of 82.9%. BLUE – Apo-SodA supplemented with $FeCl_2$ prior to refolding, total percentage inhibition of 5.9%. RED – Apo-SodA, no added co-factor, total percentage inhibition of 3%.

4.8 Crystallisation trials of *B. burgdorferi* SodA

Refolded *B. burgdorferi* SodA was concentrated to 10 and 15 mg/ml in 0.3M NaCl, 50mM tris at pH8 using 10,000K MWCO centrifugal concentrator using manufacturer instructions and centrifuged at 13,000rpm in a bench top centrifuge for 20 minutes prior to crystal screening (section 2.2.9). Hanging drops trays were set up in duplicate and by hand using 24 well plates at a ratio of 1:1 whereby the drop consisted of 1µl of protein solution with the same amount of well solution. All crystal screens were made in house from stock solutions and are summarised in table 4.4.

Trays were observed immediately after set up using a light microscope and stored at both 20 and 10^{o}C. Crystal trays were further checked on day 2, 4, 7 and weekly after this period for any changes.

Crystal Screen	Manufacturer	Formulation
Clear strategy crystal screen 1 (CSSI)	Molecular dimensions	In house
Clear strategy crystal screen 2 (CSSII)	Molecular dimensions	In house
Wizard Screen 1	Rigaku	In house
Wizard Screen 2	Rigaku	In house
Wizard Screen 3	Rigaku	In house
Wizard Screen 4	Rigaku	In house

Table 4.4. Crystallisation screens used for SOD crystallisation attempts. All crystal screens were made in house using formulation recipes from Rigaku.

Screening using recombinant SodA failed to produce any crystals with the exception of some salt crystals which were apparent in conditions containing ammonium sulphate and were identified using Izit dye and by crushing.

4.9 Homology Modelling: *Borrelia burgdorferi* superoxide dismutase

Homology models for *B. burgdorferi* SodA were generated using PHYRE2 (Kelley, 2015). Throughout the project two homology modelling methods were utilised these included PHYRE2 (Kelley, 2015) and Chimera (Pettersen et al., 2004) with Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000). The program PHYRE2 (Kelley, 2015) was the preferred method and can be considered the most automated approach. When templates with a high degree of sequence similarity can be found PHYRE2 (Kelley, 2015) can rapidly produce homology models and predict ligand binding and active site structures. This tool was highly effective for SOD modelling as the enzyme is highly conserved in contrast Borrelia OMPs shared no level of sequence similarity to other known proteins and Chimera (Pettersen et al., 2004) and Modeller (UCSF – Webb & Sali, 2014, Martin-Renom et al., 2000, Sali & Blundell, 1993, Fiser et al., 2000) provided a powerful alternative when alignment editing was required.

Protein BLAST (Altschul et al., 1990) searches identified a manganese superoxide dismutase from *T. thermophilus* with the highest sequence similarity of 51%. The structure of this protein (PDB 3MDS – Ludwig et al., 1991) along with the next four highest scoring proteins were selected by PHYRE2 (Kelley, 2015) for homology modelling and the generated structures shown in figure 4.12 and full template information shown in table 4.5.

Model	Template	Sequence similarity	Confidence	Coverage
1	Mn-SOD - *T. thermophilus* (3MDS)	51%	100%	99%
2	Mn-SOD - Nostoc sp. (1GV3)	42%	100%	99%
3	Mn-SodA-2 – *Bacillus anthracis* (1XRE)	45%	100%	99%
4	Mn-SOD – *Bacillus subtilis* (2RCV)	47%	100%	99%
5	Fe-SOD – *Clostridium difficile* (3TJT)	43%	100%	99%

Table 4.5. Summary PHYRE2 data for *B. burgdorferi* SodA homology modelling.

The model ranked highest by PHYRE2 (Kelley, 2015) was based upon a known Mn-SOD from *T. thermophilus* which shared a significant level of sequence similarity. As this model was calculated with 100% confidence the generated model was submitted to the 3D Ligand Site (Wass et al., 2010) with the aim of predicting potential binding sites. Predicted binding sites are shown in figure 4.14 with the server identifying His26, His81, Asp164 and His168 as possible residues within the binding site. This agreed particularly well with previous modelling work undertaken by Aguirre et al., 2013 previously discussed in figure 1.26 of section 1.6.5.

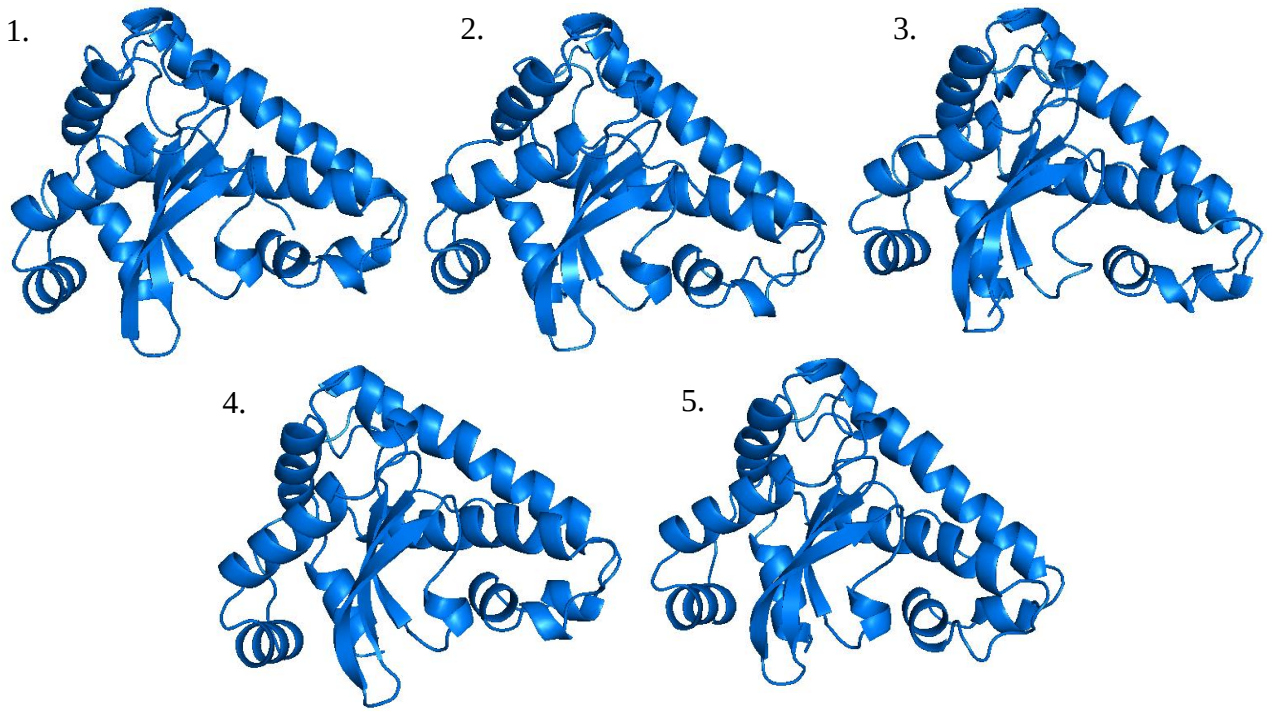


Figure 4.12. PHYRE2 homology models for *B. burgdorferi* SodA. The top five ranking homology models for *B. burgdorferi* SodA based on five independent templates. 1- *B. burgdorferi* SodA homology model using a Mn-SOD from *T. thermophilus* (PDB code - 3MDS – Ludwig et al., 1991) as a template. 2- Homology model generated using a Mn-SOD from *Nostoc sp*. (PDB code – 1GV3 – Atzenhofer et al., 2002). 3 – Generated using Mn-SodA-2 from *Bacillus anthracis* (PDB code – 1XRE – Boucher et al., 2005) as a template. 4 – Mn-SOD from *Bacillus subtilis* (PDB code – 2RCV – Liu et al., 2007) used as a template. 5 – A homology model generated using and Fe-SOD from *Clostridium difficile* (PDB code – 3TJT – Li et al., 2015) as a template.

Figure 4.13. PHYRE2 secondary structure and disorder prediction of *B. burgdorferi* SodA. Output data generated by PHYRE2 (Kelley, 2015) summarising the predicted secondary structure of *B. burgdorferi* SodA.

The PHYRE2 (Kelley, 2015) prediction was further supported following submission of the amino acid sequence to RaptorX (Källberg et al., 2012) a web tool for protein structure and function prediction. Alike PHYRE2 (Kelley, 2015) a tertiary structure for *B. burgdorferi* SodA was generated using the highest ranking template based upon the *B. burgdorferi* amino acid sequence. Alike PHYRE2 (Kelley, 2015) this was deemed to be a Mn SOD from *T. thermophilus* (3MDS - Ludwig et al., 1991) following generation of a homology model as shown in figure 4.14. This model was used for ligand binding prediction. Mn was identified as the ligand with His24, His81, Asp164 and His168 predicted to be participating within the binding site, full parameters and scoring is shown in table 4.6.

Models built using RaptorX (Källberg et al., 2012) are assessed for their reliability using three major parameters. These include the p-value, the score and the uGDT (GDT). The p-value represents the likelihood of the generated model being worse than the best set of randomly generated models for the

protein or domain in question. Therefore the lower the p-value the higher the quality of the computed model. As a general rule the p-value of a good model should close to 10^-3 for alpha proteins and less than 10^-4 for beta proteins. The second value is known as the score and is based upon the alignment scoring which can fall between 0 and the sequence length with the value of 0 being the lowest score. In some cases the score can exceed the sequence length due to error estimations.

The final value in order to assess the model is the uGDT (GDT), the uGDT is the unnormalised global distance test which is defined as, 1*N(1)+0.75*N(2)+0.5*N(4)+0.25*N(8)
where N(x) is the number of residues with estimated modelling error (in Å) smaller than x. From this, GDT is calculated as uGDT divided by the protein length and multiplied by a 100.

The uGDT(GDT) measures the absolute model quality. For a protein with >100 residues, uGDT>50 is a good indicator. For a protein with <100 residues, GDT>50 is a good indicator. If a model has good uGDT (>50) but bad GDT (<50), it indicates that only a small portion of the model may be good.

To judge the quality of the generated model all three values are used. A model with a good o-value and uGDT is likely to be of high quality in contrast a model with a good p-value but poor uGDT may be of low quality even though the p-value indicates it is better than the set of randomly generated models. Models with a low p-value but a good uGDT value may of fair quality although they may not be better than the set of randomly generated models. All three of the values exhibit a bias towards sequence or domain length. The longer the amino acid sequence submitted to RaptorX (Källberg et al., 2012) the more likely the software can build a better model with a better score, uGDT and p-value. This is somewhat rectified by taking the GDT into account when evaluating the quality of the generated model. However, other values may be useful in order to counteract the sequence length bias.

The model generated by RaptorX (Källberg et al., 2012) for *B. burgdorferi* SodA was based upon the same *T. thermophilus* template (identity 51%) as selected by PHYRE2 (Kelley, 2015) and the model scored highly with a score of 216 a p-value of 1.92e-10 and an overall uGDT (GDT) 180 (88). This high quality model was used for ligand binding searching with the predicted active site shown in figure 4.14.

Highest ranking template	Mn-SOD from *T. thermophilus* (3MDS)
Predicted number of domains	1
p-value	1.92e-10
Overall uGDT(GDT)	180 (88)
uSeqId(SeqId)	101 (50)
Score	216*
Number of residues modelled	203 (100%)
Number of positions predicted as disordered	0 (0%)
Secondary structure prediction	57% α-helix, 7% β-sheet, 34% loops
Solvent access prediction	29% exposed, 31% medium, 39% buried

Table 4.6. RaptorX summary of parameters and estimation values. All values were calculated by RaptorX (Källberg et al., 2012) structure prediction and Raptor X (Källberg et al., 2012) binding prediction. *The score value exceed the total sequence length due to error estimations.

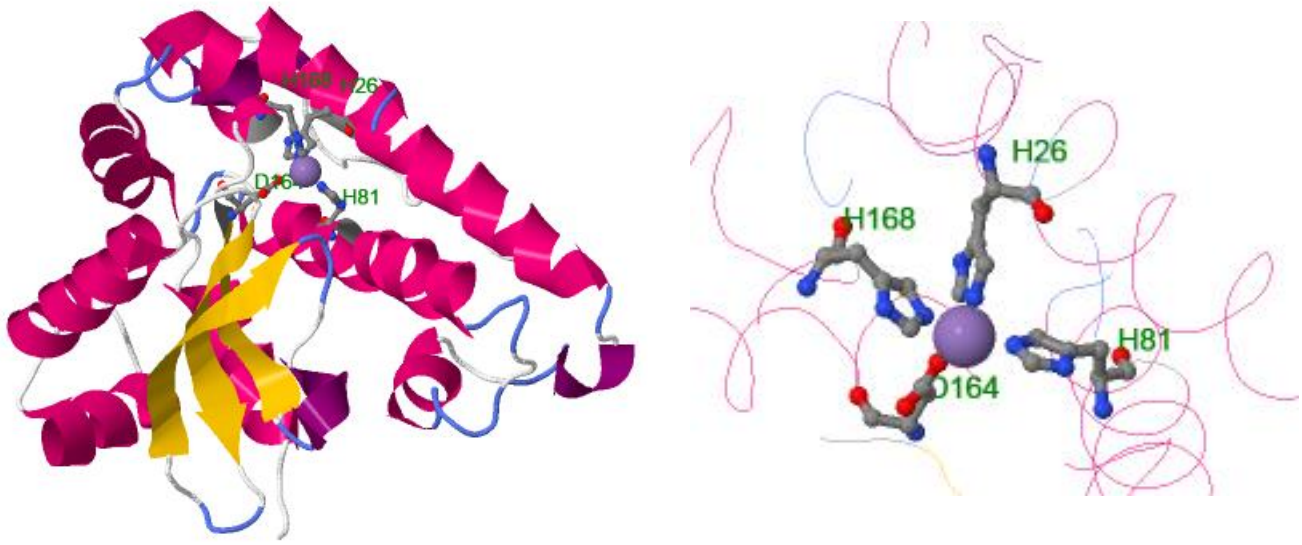


Figure 4.14. Predicted active site for *B. burgdorferi* SodA. Amino acids involved within the active site were identified using RaptorX (Källberg et al., 2012) from an initial homology model generated using a Mn-SOD from T. thermophilus (1MNG – Lah et al., 1995). LEFT – Cartoon representation of the homology model generated by RaptorX (Källberg et al., 2012) with the four predicted amino acids involved within the ligand binding site. RIGHT – A closer look at the four predicted amino acids and the Mn co-factor.

4.10 Discussion

As with many *Borrelia* genes, the *B. burgdorferi* SodA contains numerous rare codons. In the hope of improving expression by *E. coli* the gene was commercially codon optimised for use within an *E. coli* expression system. High expression levels were achieved under IPTG control and using autoinduction methods. However, SodA was observed to form insoluble inclusion bodies under all conditions tested (section 4.2). This negatively impacted the final yield as SodA required extraction and refolding from these inclusion bodies. A number of SODs from both prokaryotes (Whittaker & Whittaker, 2000, Love et al., 2002, Brennan et al., 2015) and eukaryotes (Shi et al., 1999, Sujiwattanarat et al., 2016) have been demonstrated to be produced as active soluble proteins following expression within *E. coli.* However, this was not seen for *B. burgdorferi* SodA suggesting an alternative expression system such as *Pichia pastoris* maybe useful for any future work. Interestingly around 30% of native *E. coli* proteins would also not be expressed as soluble polypeptides when overexpressed in *E. coli* (Vincentelli et al., 2003) highlighting the challenges faced when producing *Borrelia* proteins within this expression system.

Refolding yields were fair with 1L of *E. coli* expression culture often delivering 10-20 mg of recombinant SodA following gel filtration. These yields were not observed during the production of apo-SodA where the amount of final product dropped over ten-fold indicating that the metal ion co-factor plays a major role in stability. Of the small amount of apo-protein that could be obtained this sample was confirmed as folded by circular dichroism. It was determined that less apo-SodA would precipitate if metal ions were stripped from the sample prior to refolding rather than taking a sample of folded SodA and removing bound metal ions either by dialysis or by chromatography.

Both in-house and synchrotron SAXS experiments were problematic (section 4.6). Samples of Mn-SodA subjected to data collection on a Bruker nanostar exhibited radiation damage and progressive aggregation and precipitation during the exposure time. This occurred as early as the first frame of 2400 seconds making data analysis impossible. Samples of Mn-SodA taken to DLS were of good quality and enzymatically active. However, the samples precipitated instantly upon X-ray exposure. Literature searches have failed to identify any similar problems for SODs or for any other metalloprotein. However, the observed behaviour of SodA in response to X-ray exposure suggests that oxidation of the metal ion may have induced instability leading to aggregation.

Previous studies have identified that *B. burgdorferi* SodA is likely a Mn-SOD in contrast to its initial characterisation as a Fe-SOD. However, this is yet to be confirmed using a pure recombinant sample.

This study confirms that *B. burgdorferi* SodA is indeed active in the presence of manganese and has limited activity when it is stripped of this ion or when supplemented with iron (section 4.7). The low level of activity observed for Apo-SodA and Fe-SodA is likely due to residual Mn ions within the sample. For the first time it has been demonstrated that *B. burgdorferi* SodA does not exhibit cambialistic behaviour to any significant degree. The limitations of the enzymatic study is the acceptance that each SodA sample contained the desired metal ion. Atomic absorption spectroscopy could be useful in order to confirm presence of Fe and Mn within the protein sample (Del Vechhio et al., 2009).

Early crystallisation attempts of *B. burgdorferi* SodA were promising. The protein appeared fairly stable and clear drops were observed across a number of conditions with minimal amorphous aggregation observed. However, optimisation of these conditions failed to yield crystals (section 4.8). Optimisation revolved around increasing the protein concentration and increasing the precipitant concentration within each condition. Crystallisation experiments were carried out both under reducing and non-reducing conditions and where possible the reducing agents DTT and β-mercaptoethanol was supplemented with fresh stock every few days to maintain the reducing environment. Future work should aim to produce *B. burgdorferi* SodA mutants devoid of the cysteine residue at position 136 in order to improve stability for crystallisation and to conclude whether the dimer formation is physiologically relevant and required for enzymatic activity.

As previous discussed *B. burgdorferi* SodA shares a significant level of sequence similarity to a Mn-SOD from *Thermus sp*. (section 4.1) this brought to question whether this would be seen for other *Borrelia* proteins. In order to deduce whether this similarity was an exception or whether the similarity between the two species extended further, a selection of housekeeping and other important cellular *Borrelia* proteins were selected and subjected to protein BLAST (Altschul et al., 1990) searching. The full list and results can be seen within the appendix 3. This small investigation highlighted a number of *Borrelia* proteins with sequence similarity to other extremophiles rather that other spirochaetes such as *Leptospira* and *Treponema*. Further work is needed in order to decipher this shared identity with the possibility that some *Borrelia* proteins may be of commercial interest should they retain function under extreme conditions.

Unlike many other micro-organisms, *B. burgdorferi* appears to have evolved without any known requirement for iron and several studies have demonstrated that the spirochaete accumulates very high levels of cellular manganese when compared to organisms such as *E. coli* and *S. cerevisiae*

(Aguirre et al., 2013). This high manganese environment may serve an important purpose for this unique micro-organism. Firstly the absence of cellular iron and the lack of iron requiring enzymes encoded within the *Borrelia* genome suggests that manganese may be much more widely used as a co-factor for the spirochaetes metalloproteins. When *B. burgdorferi* SodA is expressed within *S. cerevisiae* the enzyme is only activated once cellular manganese levels exceeds the concentration of mitochondrial iron. These levels of manganese simulate the normal conditions within *Borrelia* (Aguirre et al., 2013).

The high manganese conditions observed within the spirochaete may also be considered a unique adaption which provides the micro-organism with an exceptional response to metal starvation. The effects of ROS upon pathogenic bacteria has been of high interest for a number of years. However, much of this work has revolved around the study of the effects of ROS upon bacteria that require iron and possess a TCA cycle (Shimizu, 2013, Rosato et al., 2014, Thomas et al., 2013). The overall effects of ROS are directly related to a micro-organisms cellular iron, due to Fenton chemistry. As *Borrelia* have no need for iron the organism presents an interesting platform to further our understanding of the effects of oxidative stress in a system devoid of iron. As the destruction of Fe-metalloproteins or damage to DNA (Troxell et al., 2012) does not appear to be the primary target of ROS within *Borrelia* the presence of the spirochaetes SodA enzyme questions its purpose. Recent work has identified that the SodA protein may offer protection from ROS mediated damage of membrane polyunsaturated fatty acids. However, if this was the primary purpose of the SodA enzyme should we not expect this protein to be membrane bound or localised within the periplasm as a result of selective pressure? (Troxell et al., 2012). As most Mn-SODs are considered intracellular proteins (Gregory et al., 1973), and it is currently accepted that *B. burgdorferi* SodA is a cytoplasmic enzyme, it is likely that the SodA protein does not only have roles within protecting the membrane. Recent work has identified that the SodA gene is essential for the spirochaetes resistance to streptonigrin (Troxell et al., 2012) which is curious as streptonigrin toxicity is dependent on iron (Yeowell & White, 1982). Streptonigrin is naturally produced by *Streptomyces flocculus* and is considered to exhibit both anti-tumour and anti-bacterial activity (Rao & Cullen, 1960). It was observed from as early as the 1980's that the activity of streptonigrin was hindered should a bacterium become unable to uptake iron (Yeowell & White, 1982). The overall effect of streptonigrin within the cell is the promotion of ROS (Hassan & Fridovich, 1979) and NADH is thought to aid the redox cycling of streptonigrin (White & White, 1968, Troxell et al., 2012). The sensitivity of *B. burgdorferi* to this compound remains poorly understood and the presence of either manganese or copper could be involved within this reaction

(Troxell et al., 2012). More importantly this sensitivity highlights that *Borrelia* may be susceptible to damage by other redox cycling compounds which the SodA enzyme offers protection from.

In conclusion *B. burgdorferi* SodA appears as an essential and unique protein in a family of normally well conserved enzymes. It is likely that SodA plays other roles other than protection to polyunsaturated fatty acids within the membrane which will at some point explain its intracellular localisation. The *Borrelia* system also provides an interesting and unique platform for the study of ROS damage in a system where iron is not required and may be of importance for improving our understanding of redox cycling drugs.

Chapter 5: References

Abdul-Gader, A., Miles, A.J., Wallace, B.A. (2011) A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics* **27**, 1630-1636.

Adams, R., Worth, C., Guenther, S., Dunkel, M., Lehmann, R. and Preissner, R. (2011) 'Binding sites in membrane proteins--diversity, druggability and prospects', *European journal of cell biology.*, 91(4), pp. 326–39.

Adeolu, M. and Gupta, R. S. (2014) 'A phylogenomic and molecular marker based proposal for the division of the genus Borrelia into two genera: the emended genus Borrelia containing only the members of the relapsing fever Borrelia, and the genus Borreliella gen. nov. containing the members of the Lyme disease Borrelia (Borrelia burgdorferi sensu lato complex)' Antonie Van Leeuwenhoek, 105(6): 1049-1072.

Aguero-Rosenfeld, M. E., Nowakowski, J., Bittker, S., Cooper, D., Nadelman, R. B. and Wormser, G. P. (1996) 'Evolution of the serologic response to Borrelia burgdorferi in treated patients with culture-confirmed erythema migrans', available at http://jcm.asm.org/content/34/1/1.long, accessed 22 March 2016.

Aguero-Rosenfeld, M. E., Nowakowski, J., McKenna, D. F., Carbonaro, C. A. and Wormser, G. P. (1993) 'Serodiagnosis in early Lyme disease', *Journal of Clinical Microbiology*, 31 (12): 3090–3095.

Aguilar, C., Thomas, P., Mills, A., Moss, D. and Palmer, R. (1992) 'Newly observed binding mode in pancreatic ribonuclease', Journal of molecular biology., 224(1), pp. 265–7.

Aguirre, J. D., Clark, H. M., McIlvin, M., Vazquez, C., Palmere, S. L., Grab, D. J., Seshu, J., Hart, P. J., Saito, M. and Culotta, V. C. (2013) 'A manganese-rich environment supports Superoxide Dismutase activity in a Lyme disease Pathogen, Borrelia burgdorferi', *Journal of Biological Chemistry*, 288 (12): 8468–8478.

Aherne, M., Lyons, J. A., & Caffrey, M. (2012). A fast, simple and robust protocol for growing crystals in the lipidic cubic phase. *Journal of Applied Crystallography*, *45*(Pt 6), 1330–1333.

Ahn, V.E., Lo, E.I., Engel, C.K., Chen, L., Hwang, P.M., Kay, L.E., Bishop, R.E. and Privé, G.G. (2004) 'A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin', *EMBO*, 23(15), pp. 2931–2941.

Ai, C., Wen, Y., Zhang, Y., Wang, S., Qiu, Q., Shi, Z., Li, D., Chen, D., Liu, X. and Zhao, J. (1988) 'Clinical manifestations and epidemiological characteristics of Lyme disease in Hailin county, Heilongjiang province, china', *Annals of the New York Academy of Sciences.*, 539 302–13.

Alban, P., Johnson, P. and Nelson (2000) 'Serum-starvation-induced changes in protein synthesis and morphology of Borrelia burgdorferi', *Microbiology*, 146. 119–27.

Alitalo, A., Meri, T., Ramo, L., Jokiranta, T. S., Heikkila, T., Seppala, I. J. T., Oksi, J., Viljanen, M. and Meri, S. (2001) 'Complement evasion by Borrelia burgdorferi: Serum-resistant strains promote C3b inactivation', *Infection and Immunity*, 69 (6): 3685–3691.

Allen, J.P. and Feher, G. (1984) 'Crystallization of reaction center from *Rhodopseudomonas sphaeroides*: Preliminary characterization', *Biophysics*, 81: 4795-4799.

Alloing, G., de Philip, P. and Claverys, J.P. (1994) 'Three highly homologous membrane-bound lipoproteins participate in Oligopeptide transport by the Ami System of the gram-positive Streptococcus pneumoniae', *Journal of Molecular Biology*, 241 (1): 44–58.

Altpetera, E., Zimmermanna, H., Oberreicha, J., Péterb, O. and Dvořákc, C. (2013) 'Tick related diseases in Switzerland, 2008 to 2011.', *Swiss Medical Weekly*, 143.

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.

Anguita, J., Ramamoorthi, N., Hovius, J., Das, S., Thomas, V., Persinski, R., Conze, D., Askenase, P., Rincón, M., Kantor, F. and Fikrig, E. (2002) 'Salp15, an ixodes scapularis salivary protein, inhibits CD4(+) T cell activation', *Immunity.*, 16(6): 849–859.

Antonara, S., Chafel, R. M., LaFrance, M. and Coburn, J. (2007) 'Borrelia burgdorferi adhesins identified using in vivo phage display', *Molecular Microbiology*, 66 (1): 262–276.

Arai, S. and Hirai, M. (1999) 'Reversibility and hierarchy of thermal transition of hen egg-white lysozyme studied by small-angle x-ray scattering', *Biophysical journal.*, 76(4): 2192–2197.

Aree, T., Hoier, H., Schulz, B., Reck, G. and Saenger, W. (2000) 'Novel type of Thermostable channel Clathrate Hydrate formed by Heptakis(2, 6-di-O-methyl)-β-cyclodextrin-15 H_2O-A paradigm of the hydrophobic effect', *Angew Chem Int Ed Engl*, 39 (5): 897–899.

Arora, A., Rinehart, D., Szabo, G. and Tamm, L.K. (2000) 'Refolded outer membrane protein A of Escherichia coliForms ion channels with Two Conductance states in planar lipid Bilayers', Journal of Biological Chemistry, 275(3), pp. 1594–1600.

Arumugam TV, Magnus T, Woodruff TM, Proctor LM, Shiels IA, Taylor SM. (2006) 'Complement mediators in ischemia-reperfusion injury'. *Clin Chim Acta*. 374(1-2):33–45.

Asada K., Yoshikawa K., Takahashi M., Maeda Y., Enmanji K. (1975) Superoxide dismutases from a blue-green alga, *Plectonema boryanum*. J. Biol. Chem. 250, 2801–2807.

Åsbrink, E. and Hovmark, A. (1988) 'Early and late Cutaneous manifestations in Ixodes-borne Borreliosis (Erythema Migrans Borreliosis, Lyme Borreliosis)', *Annals of the New York Academy of Sciences*, 539 (1 Lyme Disease): 4–15.

Asbrink, E., Olsson, I. and Hovmark, A. (1986) 'Erythema chronicum migrans Afzelius in Sweden. A study on 231 patients', *Zentralblatt für Bakteriologie, Mikrobiologie, und Hygiene. Series A, Medical microbiology, infectious diseases, virology, parasitology*. 263: 229–236.

Atzenhofer, W., Regelsberger, G., Jacob, U., Peschek, G., Furtmüller, P., Huber, R. and Obinger, C. (2002) 'The 2.0A resolution structure of the catalytic portion of a cyanobacterial membrane-bound manganese superoxide dismutase', Journal of molecular biology., 321(3), pp. 479–89.

Babb, K., El-Hage, N., Miller, J. C., Carroll, J. A. and Stevenson, B. (2001) 'Distinct regulatory pathways control expression of Borrelia burgdorferi infection-associated OspC and Erp surface proteins', *Infection and Immunity*, 69 (6): 4146–4153.

Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L. and Sankaran, K. (2006) 'A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins', *Journal of Bacteriology*, 188 (8): 2761–2773.

Barbas, C.F., Burton, D.R., Scott, J.K. and Silverman, G.J. (2007) 'Quantitation of DNA and RNA', Cold Spring Harbor Protocols, (11), p. 47.

Barbour, A. (2000) 'Antigenic variation in vector-borne Pathogens', *Emerging Infectious Diseases*, 6 (5): 449–457.

Barbour, A. G. (1984) 'Isolation and cultivation of Lyme disease spirochetes', *Yale J Biol Med*, 57 (4): 521–525.

Barbour, A. G. and Hayes, S. F. (1986) 'Biology of Borrelia species', *Microbiol Rev*, 50 (4): 381–400.

Barbour, A. G. and Zückert, W. R. (1997) 'Genome sequencing: New tricks of tick-borne pathogen', *Nature*, 390 (6660): 553–554.

Barbour, A. and Garon, C. (1987) 'Linear plasmids of the bacterium Borrelia burgdorferi have covalently closed ends', *Science*, 237 (4813): 409–411.

Barthold, S. W. (1996) 'Globalisation of Lyme borreliosis', *The Lancet*, 348 (9042): 1603–1604.

Bartra, S., Gong, X., Lorica, C., Jain, C., Nair, M., Schifferli, D., Qian, L., Li, Z., Plano, G. and Schesser, K. (2011) 'The outer membrane protein A (OmpA) of Yersinia pestis promotes intracellular survival and virulence in mice', *Microbial pathogenesis.*, 52 (1): 41–46.

Beaurepaire, C. and Chaconas, G. (2005) 'Mapping of essential replication functions of the linear plasmid lp17 of B. Burgdorferi by targeted deletion walking', *Molecular Microbiology*, 57 (1): 132–142.

Becker, M., Bunikis, J., Lade, B.D., Dunn, J.J., Barbour, A.G. and Lawson, C.L. (2005) 'Structural investigation of Borrelia burgdorferi OspB, a BactericidalFab target', Journal of Biological Chemistry, 280(17), pp. 17363–17370.

Beermann, C., Lochnit, G., Geyer, R., Groscurth, P. and Filgueira, L. (2000) 'The lipid component of lipoproteins from Borrelia burgdorferi: Structural analysis, Antigenicity, and presentation via human Dendritic cells', *Biochemical and Biophysical Research Communications*, 267 (3): 897–905.

Beher, M., Schnaitman, C. and Pugsley, A. (1980) 'Major heat-modifiable outer membrane protein in gram-negative bacteria: Comparison with the ompA protein of Escherichia coli', *Journal of bacteriology.*, 143 (2): 906–913.

Behera, A., Durand, E., Cugini, C., Antonara, S., Bourassa, L., Hildebrand, E., Hu, L. and Coburn, J. (2007) 'Borrelia burgdorferi BBB07 interaction with integrin alpha3beta1 stimulates production of pro-inflammatory mediators in primary human chondrocytes', *Cellular microbiology.*, 10 (2): 320–331.

Bell, L., Alpert, G., Campos, J. and Plotkin, S. (1985) 'Routine quantitative blood cultures in children with Haemophilus influenzae or Streptococcus pneumoniae bacteremia', *Pediatrics.*, 76 (6): 901–907.

Ben-Menachem, G., Kubler-Kielb, J., Coxon, B., Yergey, A. and Schneerson, R. (2003) 'A newly discovered cholesteryl galactoside from Borrelia burgdorferi', *Proceedings of the National Academy of Sciences of the United States of America.*, 100 (13): 7913–8.

Benke, T., Gasse, T., Hittmair-Delazer, M. and Schmutzhard, E. (1995) 'Lyme encephalopathy: Long-term neuropsychological deficits years after acute neuroborreliosis', *Acta neurologica Scandinavica.*, 91 (5): 353–7.

Bennet, L. and Berglund, J. (2002) 'Reinfection with Lyme borreliosis: A retrospective follow-up study in southern Sweden', *Scandinavian journal of infectious diseases.*, 34 (3): 183–6.

Benov L. T., Fridovich I. (1994) *Escherichia coli* expresses a copper- and zinc-containing superoxide dismutase. J. Biol. Chem. 269, 25310–25314.

Benvenuti, M. and Mangani, S. (2007) 'Crystallization of soluble proteins in vapor diffusion for x-ray crystallography: Abstract: Nature protocols', Nature Protocols, 2(7), pp. 1633–1651.

Bercic, R., Slavec, B., Lavric, M., Narat, M., Bidovec, A., Dovc, P. and Bencina, D. (2007) 'Identification of major immunogenic proteins of Mycoplasma synoviae isolates', *Veterinary microbiology.*, 127: 147–54.

Berglund, J., Eitrem, R., Ornstein, K., Lindberg, A., Ringér, A., Elmrud, H., Carlsson, M., Runehagen, A., Svanborg, C. and Norrby, R. (1995) 'An epidemiologic study of Lyme disease in southern Sweden', *The New England journal of medicine.*, 333 (20): 1319–27.

Berglund, J., Stjernberg, L., Ornstein, K., Tykesson-Joelsson, K. and Walter, H. (2002) '5-y follow-up study of patients with neuroborreliosis', *Scandinavian journal of infectious diseases.*, 34 (6): 421–5.

Berglund, J., Stjernberg, L., Ornstein, K., Tykesson-Joelsson, K. and Walter, H. (2002) '5-y follow-up study of patients with neuroborreliosis', *Scandinavian journal of infectious diseases.*, 34 (6): 421–5.

Bergström, S. and Zückert, W.R. (2010) 'Structure, function and biogenesis of the Borrelia cell envelope',*Borrelia: molecular biology, host interaction and pathogenesis* 139–166.

Berman, P. and Banker, B. (1966) 'Neonatal meningitis. A clinical and pathological study of 29 cases', *Pediatrics.*, 38 (1): 6–24.

Berndtson, K. (2013) 'Review of evidence for immune evasion and persistent infection in Lyme disease', *International journal of general medicine.*, 6: 291–306.

Bestor, A., Rego, R., Tilly, K. and Rosa, P. (2012) 'Competitive advantage of Borrelia burgdorferi with outer surface protein BBA03 during tick-mediated infection of the mammalian host', *Infection and immunity.*, 80 (10): 3501–11.

Bettridge, J., Renard, M., Zhao, F., Bown, K. and Birtles, R. (2013) 'Distribution of Borrelia burgdorferi sensu lato in Ixodes ricinus populations across central Britain', *Vector borne and zoonotic diseases (Larchmont, N.Y.).*, 13 (3): 139–46.

Bhattacharya, S., Ploplis, V. and Castellino, F. (2012) 'Bacterial plasminogen receptors utilize host plasminogen system for effective invasion and dissemination', *Journal of biomedicine & biotechnology.*, 2012, Article ID 482096.

Bhide M.R., Escudero R., Camafeita E., Gil H., Jado I., Anda P. (2009) 'Complement factor H binding by different Lyme disease and relapsing fever Borrelia in animals and human', *BMC Res. Notes*.2:134.

Blackmore, T., Hellwage, J., Sadlon, T., Higgs, N., Zipfel, P., Ward, H. and Gordon, D. (1998) 'Identification of the second heparin-binding domain in human complement factor H', *Journal of immunology (Baltimore, Md. : 1950).*, 160(7), pp. 3342–8.

Bolduc, G.R., Bouchet, V., Jiang, R.-Z., Geisselsoder, J., Truong-Bolduc, Q.C., Rice, P.A., Pelton, S.I. and Goldstein, R. (2000) 'Variability of outer membrane protein P1 and its evaluation as a vaccine candidate against experimental Otitis media due to Nontypeable Haemophilus influenzae: An unambiguous, multifaceted approach', *Infection and Immunity*, 68 (8): 4505–4517.

Borchers, A., Keen, C., Huntley, A. and Gershwin, M. (2014) 'Lyme disease: A rigorous review of diagnostic criteria and treatment', *Journal of autoimmunity.*, 57 82–115.

Bordier, C. (1981) 'Phase separation of integral membrane proteins in Triton X-114 solution', *The Journal of biological chemistry.*, 256(4), pp. 1604–7.

Borg, R., Dotevall, L., Hagberg, L., Maraspin, V., Lotric-Furlan, S., Cimperman, J. and Strle, F. (2005) 'Intravenous ceftriaxone compared with oral doxycycline for the treatment of Lyme neuroborreliosis', *Scandinavian journal of infectious diseases.*, 37 449–54.

Boucher, I.W., Kalliomaa, A.K., Levdikov, V.M., Blagova, E.V., Fogg, M.J., Brannigan, J.A., Wilson, K.S. and Wilkinson, A.J. (2005) 'Structures of two superoxide dismutases from Bacillus anthracis reveal a novel active centre', Acta Crystallogr Sect F Struct Biol Cryst Commun, 61(Pt 7), pp. 621–624.

Bouchon, B., Klein, M., Bischoff, R., Dorsselaer, V. and Roitsch, C. (1997) 'Analysis of the lipidated recombinant outer surface protein A from Borrelia burgdorferi by mass spectrometry', *Analytical biochemistry.*, 246 (1): 52–61.

Bradford, M.M. (1976), 'Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding', *Anal. Biochem.*, **72**: 248–254.

Breitner-Ruddock, S., Würzner, R., Schulze, J. and Brade, V. (1997) 'Heterogeneity in the complement-dependent bacteriolysis within the species of Borrelia burgdorferi', *Medical microbiology and immunology.*, 185 (4): 253–60.

Bremell, D., Säll, C., Gisslén, M. and Hagberg, L. (2011) 'Lyme neuroborreliosis in HIV-1 positive men successfully treated with oral doxycycline: A case series and literature review', *Journal of Medical Case Reports*, 5 (1): 465.

Bremer, H. (1982) 'Variation of generation times in Escherichia coli populations: Its cause and implications', *Journal of general microbiology.*, 128 (12): 2865–76.

Brennan, R.E., Kiss, K., Baalman, R. and Samuel, J.E. (2015) 'Cloning, expression, and characterization of a Coxiella burnetii cu/Zn Superoxide dismutase', *BMC Microbiology*, 15(1).

Brissette, C., Rossmann, E., Bowman, A., Cooley, A., Riley, S., Hunfeld, K., Bechtel, M., Kraiczy, P. and Stevenson, B. (2009) 'The borrelial fibronectin-binding protein RevA is an early antigen of human Lyme disease', *Clinical and vaccine immunology : CVI.*, 17 (2): 274–80.

Brisson, D., Drecktrah, D., Eggers, C. and Samuels, D. (2012) 'Genetics of Borrelia burgdorferi', *Annual review of genetics.*, 46 515–36.

British Infection Association (2011) 'The epidemiology, prevention, investigation and treatment of Lyme borreliosis in United Kingdom patients: A position statement by the British infection association British infection association', *Journal of Infection*, 62 329–338.

Brooks, C., Vuppala, S., Jett, A. and Akins (2005) 'Identification of Borrelia burgdorferi outer surface proteins', *Infection and immunity.*, 74 (1): 296–304.

Brosig, A., Nesper, J., Boos, W., Welte, W. and Diederichs, K. (2008) 'Crystal structure of a major outer membrane protein from Thermus thermophilus HB27', *Journal of molecular biology.*, 385(5), pp. 1445–55.

Brown, D. and London, E. (1999) 'Functions of lipid rafts in biological membranes', *Annual review of cell and developmental biology.*,14 111–36.

Bu, X., Yao, X., Jiao, S., Zeng, F., Liu, Y., Xiang, Y., Liang, C., Wang, Q., Wang, X., Cao, H., Yi, X., Deng, B., Liu, C., Xu, J., Zhang, L., Gao, C., Xu, Z., Zhang, M., Tan, X., Zhou, H. and Wang, Y. (2014) 'A study on the association between infectious burden and Alzheimer's disease', *European journal of neurology.*, 22 (12): 1519–25.

Bu, Z., Koide, S. and Engelman, D. (1998) 'A solution SAXS study of Borrelia burgdorferi OspA, a protein containing a single-layer beta-sheet', *Protein science : a publication of the Protein Society.*, 7 (12): 2681–3.

Buchanan, S., Smith, B., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., der, van and Deisenhofer, J. (1999) 'Crystal structure of the outer membrane active transporter FepA from Escherichia coli', *Nature structural biology.*, 6 (1): 56–63.

Buchner H. Zur Nomenklatur der schutzenden Eiweisskorper. Centr Bakteriol Parasitenk. 1891;10:699–701.

Bunikis, I., Denker, K., Östberg, Y., Andersen, C., Benz, R. and Bergström, S. (2008) 'An RND-Type Efflux system in Borrelia burgdorferi is involved in virulence and resistance to Antimicrobial compounds', *PLoS Pathogens*, 4 (2): e1000009.

Burgdorfer, W., Barbour, A., Hayes, S., Benach, J., Grunwaldt, E. and Davis, J. (1982) 'Lyme disease-a tick-borne spirochetosis?', *Science (New York, N.Y.).*, 216 (4552): 1317–9.

Bykowski, T., Woodman, M., Cooley, A., Brissette, C., Wallich, R., Brade, V., Kraiczy, P. and Stevenson, B. (2008) 'Borrelia burgdorferi complement regulator-acquiring surface proteins (BbCRASPs): Expression patterns during the mammal-tick infection cycle', *International journal of medical microbiology : IJMM.*, 298 249–56.

Caffrey, M. (2008) 'On the mechanism of membrane protein crystallization in Lipidic Mesophases †', *Crystal Growth & Design*, 8(12), pp. 4244–4254.

Caffrey, M. and Cherezov, V. (2009) 'Crystallizing membrane proteins using lipidic mesophases', *Nature protocols.*, 4(5), pp. 706–31.

Capila, I. and Linhardt, R.J. (2002) 'Heparin-Protein interactions', *Angewandte Chemie International Edition*, 41(3), pp. 390–412.

Carson, M., Johnson, D. H., McDonald, H., Brouillette, C. and Delucas, L. J. (2007) 'His-tag impact on structure', Acta Crystallography, 63(3): 295-301.

Casjens, S., Delange, M., Ley, H., Rosa, P. and Huang, W. (1995) 'Linear chromosomes of Lyme disease agent spirochetes: Genetic diversity and conservation of gene order', *Journal of bacteriology.*, 177 (10): 2769–80.

Casjens, S.R., Fraser-Liggett, C.M., Mongodin, E.F., Qiu, W.-G., Dunn, J.J., Luft, B.J. and Schutzer, S.E. (2011) 'Whole genome sequence of an unusual Borrelia burgdorferi Sensu Lato Isolate▿', *Journal of Bacteriology*, 193(6): 1489–1490.

Casjens, S., Palmer, N., Vugt, van, Huang, W., Stevenson, B., Rosa, P., Lathigra, R., Sutton, G., Peterson, J., Dodson, R., Haft, D., Hickey, E., Gwinn, M., White, O. and Fraser, C. (2000) 'A bacterial genome in flux: The twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi', *Molecular microbiology.*, 35 (3): 490–516.

Casjens, S. R., Mongodin, E. F., Qiu, W.-G., Luft, B. J., Schutzer, S. E., Gilcrease, E. B., Huang, W. M., Vujadinovic, M., Aron, J. K., Vargas, L. C., Freeman, S., Radune, D., Weidman, J. F., Dimitrov, G. I., Khouri, H. M., Sosa, J. E., Halpin, R. A., Dunn, J. J. and Fraser, C. M. (2012) 'Genome stability of Lyme disease Spirochetes: Comparative Genomics of Borrelia burgdorferi Plasmids', *PLoS ONE*, 7 (3): e33280.

Cassatt, D., Patel, N. and Ulbrandt, N. (1998) 'DbpA, but not OspA, is expressed by Borrelia burgdorferi during spirochetemia and is a target for protective antibodies', *Infection and immunity.*, 66 (11): 5379–87.
Catalina, M., Moranta, D., Regueiro, V., Llobet, E., Tomás, A., Garmendia, J. and Bengoechea, J.A. (2011) 'Klebsiella pneumoniae outer membrane protein A is required to prevent the activation of airway Epithelial cells', *Journal of Biological Chemistry*, 286 (12): 9956–9967.

CDC (2015) 'Lyme disease data tables', available at http://www.cdc.gov/lyme/stats/tables.html, accessed 29 March 2016.

Cerar, T., Ogrinc, K., Strle, F. and Ruzić-Sabljić, E. (2010) 'Humoral immune responses in patients with Lyme neuroborreliosis', *Clinical and vaccine immunology : CVI.*, 17 (4): 645–50.

Chacón, P. and Wriggers, W. (2002) ‘Multi-resolution contour-based fitting of macromolecular structures’, Journal of molecular biology., 317(3), pp. 375–84.

Chadwick, M., Westley, B. and May, F. (1997) ‘Homodimerization and hetero-oligomerization of the single-domain trefoil protein pNR-2/pS2 through cysteine 58’, The Biochemical journal., 327, pp. 117–23.

Chen, Z. and Rand, R. (1997) 'The influence of cholesterol on phospholipid membrane curvature and bending elasticity', *Biophysical journal.*, 73 (1): 267–76.

Chenoweth DE, Hugli TE. (1978) Demonstration of specific C5a receptor on intact human polymorphonuclear leukocytes. *Proc Natl Acad Sci U S A*. (8):3943–3947

Chenu, J. and Cox, J. (2009) 'Cronobacter (’Enterobacter sakazakii'): Current status and future prospects', *Letters in applied microbiology.*, 49 (2): 153–9.

Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., Choi, H.-J., Kuhn, P., Weis, W.I., Kobilka, B.K. and Stevens, R.C. (2007) ‘High resolution crystal structure of an engineered human ß2-Adrenergic G protein-coupled receptor’, Science, 318(5854), pp. 1258–1265.

Chimento, D., Mohanty, A., Kadner, R. and Wiener, M. (2003) 'Substrate-induced transmembrane signaling in the cobalamin transporter BtuB', *Nature structural biology.*, 10 (5): 394–401.

Chmelar, J., Calvo, E., Pedra, J., Francischetti, I. and Kotsyfakis, M. (2012) 'Tick salivary secretion as a source of antihemostatics', *Journal of proteomics.*, 75 (13): 3842–54.

Christova, I. and Komitova, R. (2004) 'Clinical and epidemiological features of Lyme borreliosis in Bulgaria', *Wiener klinische Wochenschrift.*, 116 42–6.

Clark, K., Leydet, B. and Hartman, S. (2013) 'Lyme borreliosis in human patients in Florida and Georgia, USA', *International journal of medical sciences.*, 10 (7): 915–31.

Cluss, R., Silverman, D. and Stafford, T. (2004) 'Extracellular secretion of the Borrelia burgdorferi Oms28 porin and Bgp, a glycosaminoglycan binding protein', *Infection and immunity.*, 72 (11): 6279–86.

Coburn, J. and Cugini, C. (2003) 'Targeted mutation of the outer membrane protein P66 disrupts attachment of the Lyme disease agent, Borrelia burgdorferi, to integrin alphavbeta3', *Proceedings of the National Academy of Sciences of the United States of America.*, 100 (12): 7301–6.

Coleman, J., Toledo, A. and Benach, J. (2015) 'Borrelia burgdorferi HtrA: Evidence for twofold proteolysis of outer membrane protein p66', *Molecular microbiology.*, 99 (1): 135–50.

Colli, C., Leinweber, B., Müllegger, R., Chott, A., Kerl, H. and Cerroni, L. (2004) 'Borrelia burgdorferi-associated lymphocytoma cutis: Clinicopathologic, immunophenotypic, and molecular study of 106 cases', *Journal of cutaneous pathology.*, 31 (3): 232–40.

Comstock, L., Fikrig, E., Shoberg, R., Flavell, R. and Thomas, D. (1993) 'A monoclonal antibody to OspA inhibits association of Borrelia burgdorferi with human endothelial cells', *Infection and immunity.*, 61 (2): 423–31.

Comstock, L. and Thomas, D. (1989) 'Penetration of endothelial cell monolayers by Borrelia burgdorferi', *Infection and immunity.*, 57 (5): 1626–8.

Compton, L.A. and Johnson, W.C., Jr. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.* **155**, 155-167.

Cook, M.J. (2015) ‘Lyme borreliosis: a review of data on transmission time after tick attachment’, *Int Journal Gen Med*, 8:1-8

Correnti, C. and Strong, R. (2012) 'Mammalian siderophores, siderophore-binding lipocalins, and the labile iron pool', *The Journal of biological chemistry.*, 287 (17): 13524–31.

Costello, C., Steere, A., Pinkerton, R. and Feder, H. (1989) 'A prospective study of tick bites in an endemic area for Lyme disease', *Connecticut medicine.*, 53 (6): 338–40.

Costello, J., Alexander, M., Greco, K., Perez-Atayde, A. and Laussen, P. (2009) 'Lyme carditis in children: Presentation, predictive factors, and clinical course', *Pediatrics.*, 123 (5).

Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N. and Rosenbusch, J. P. (1992) 'Crystal structures explain functional properties of two E. Coli porins', *Nature*, 358 (6389): 727–733.

Coyle, B., Strickland, G., Liang, Y., Peña, C., McCarter, R. and Israel, E. (1996) ‘The public health impact of Lyme disease in Maryland’, *The Journal of infectious diseases.*, 173(5), pp. 1260–2.

Craft, J., Grodzicki, R. and Steere, A. (1984) 'Antibody response in Lyme disease: Evaluation of diagnostic tests', *The Journal of infectious diseases.*, 149 (5): 789–95.

Crowley, J., Toledo, A., LaRocca, T., Coleman, J., London, E. and Benach, J. (2013) 'Lipid exchange between Borrelia burgdorferi and host cells', *PLoS pathogens.*, 9 (1).

Cuesta-Seijo, J., Neale, C., Khan, M., Moktar, J., Tran, C., Bishop, R., Pomès, R. and Privé, G. (2010) ‘PagP crystallized from SDS/cosolvent reveals the route for phospholipid access to the hydrocarbon ruler’,*Structure (London, England : 1993).*, 18(9), pp. 1210–9.

Cugini, C., Medrano, M., Schwan, T. and Coburn, J. (2003) 'Regulation of expression of the Borrelia burgdorferi beta(3)-chain integrin ligand, P66, in ticks and in culture', *Infection and immunity.*, 71 (2): 1001–7.

Cutler, S. (2006) 'Possibilities for relapsing fever reemergence', *Emerging infectious diseases.*, 12 (3): 369–74.

Dam, van, Kuiper, H., Vos, K., Widjojokusumo, A., Jongh, de, Spanjaard, L., Ramselaar, A., Kramer, M. and Dankert, J. (1993) 'Different genospecies of Borrelia burgdorferi are associated with distinct clinical manifestations of Lyme borreliosis', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.*, 17 (4): 708–17.

Dattwyler, R., Wormser, G., Rush, T., Finkel, M., Schoen, R., Grunwaldt, E., Franklin, M., Hilton, E., Bryant, G., Agger, W. and Maladorno, D. (2005) 'A comparison of two treatment regimens of ceftriaxone in late Lyme disease', *Wiener klinische Wochenschrift.*, 117 393–7.

Dawes, J. & Pepper, D.S. (1987) ‘A competitive binding assay for heparin, heparan sulphates and other sulphated polymers.’ Handbook of Experimental Pharmacology. 82: 543-559.

Den, V., Clemons, W., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. and Rapoport, T. (2003) ‘X-ray structure of a protein-conducting channel’, *Nature.*, 427(6969), pp. 36–44.

den, van, Black, P., Clemons, W. and Rapoport, T. (2004) 'Crystal structure of the long-chain fatty acid transporter FadL', *Science (New York, N.Y.).*, 304 (5676): 1506–9.

DePietropaolo, D., Powers, J., Gill, J. and Foy, A. (2006) 'Diagnosis of Lyme disease', *Delaware medical journal.*, 78 (1): 11–8.

der, van (1991) 'Lyme carditis: Clinical characteristics of 105 cases', *Scandinavian journal of infectious diseases. Supplementum.*, 77 81–4.

de Vries B, Köhl J, Leclercq WK, Wolfs TG, van Bijnen AA, Heeringa P, Buurman WA. (2003) 'Complement factor C5a mediates renal ischemia-reperfusion injury independent from neutrophils'. *J Immunol*. 1;170(7):3883–3889

Dhôte, Basse-Guerineau, A., Beaumesnil, V., Christoforov, B. and Assous, M. (2001) 'Full spectrum of clinical, serological, and epidemiological features of complicated forms of Lyme borreliosis in the Paris, France, area', *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology.*, 19 (11): 809–15.

Deitsch, K.W., Lukehart, S.A. and Stringer, J.R. (2009) 'Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens', *Nat Rev Microbio*, 7(7) 493–503.

Del Vecchio, M., Pogni, R., Baratto, M. C., Nobbs, A., Rappuoli, R., Pizza, M., & Balducci, E. (2009). Identification of an Iron-Sulfur Cluster That Modulates the Enzymatic Activity in NarE, a *Neisseria meningitidis* ADP-ribosyltransferase. *The Journal of Biological Chemistry*, *284*(48), 33040–33047.

Diamond, R. (1974) 'Real-space refinement of the structure of hen egg-white lysozyme ', Journal of Molecular Biology, 82(3), pp. 11391–5375.

Dietzman, D. E., Fischer, G. W. and Schoenknecht, F. D. (1974) 'Neonatal escherichia coli septicemia—bacterial counts in blood', *The Journal of Pediatrics*, 85 (1): 128–130.

Dorward, D. W., Fischer, E. R. and Brooks, D. M. (1997) 'Invasion and Cytopathic killing of human lymphocytes by Spirochetes causing Lyme disease', *Clinical Infectious Diseases*, 25 2–8.

Dorward, D. W., Schwan, T. G. and Garon, C. F. (1991) 'Immune capture and detection of Borrelia burgdorferi antigens in urine, blood, or tissues from infected ticks, mice, dogs, and humans', *J Clin Microbiol*, 29 (6): 1162–1170.

Drenth, J. (2007) Principles of protein x-ray crystallography. Available at: http://www.springer.com/gp/book/9780387333342 (Accessed: 18 January 2017).

Dunham-Ems, S. M., Caimano, M. J., Pal, U., Wolgemuth, C. W., Eggers, C. H., Balic, A. and Radolf, J. D. (2009) 'Live imaging reveals a biphasic mode of dissemination of Borrelia burgdorferi within ticks', *J Clin Invest.*, 119 (12): 3652–3665.

Dunn, J., Buchstein, S., Butler, L., Fisenne, S., Polin, D., Lade, B. and Luft, B. (1994) 'Complete nucleotide sequence of a circular plasmid from the Lyme disease spirochete, Borrelia burgdorferi', *Journal of bacteriology.*, 176 (9): 2706–17.

Dunn, J. P., Kenedy, M. R., Iqbal, H. and Akins, D. R. (2015) 'BMC microbiology', *BMC Microbiology*, 15 (1): 1.

Dutzler, R., Rummel, G., Albertí, S., Hernández-Allés, S., Phale, P., Rosenbusch, J., Benedí, V. and Schirmer, T. (1999) 'Crystal structure and functional characterization of OmpK36, the osmoporin of Klebsiella pneumoniae', *Structure (London, England : 1993).*, 7 (4): 425–34.

Dworkin, M.S., Schwan, T.G., Anderson, D.E. and Borchardt, S.M. (2008) 'Tick-borne Relapsing fever' *Infect Dis Clin North Am.* 22 (3): 449–468.

Dyer, A. (2013) *Identification and Structural Characterisation of Novel Outer Membrane Proteins in B. burgdorferi s.l.* Thesis thesis. University of Huddersfield.

Dyer, A., Brown, G., Stejskal, L., Laity, P.R. and Bingham, R.J. (2015) 'The Borrelia afzelii outer membrane protein BAPKO_0422 binds human factor-h and is predicted to form a membrane-spanning β-barrel',*Bioscience Reports*, 35(4).

Ebnet, K., Brown, K. D., Siebenlist, U. K., Simon, M. M. and Shaw, S. (1997) 'Borrelia burgdorferi activates nuclear factor-kappa B and is a potent inducer of chemokine and adhesion molecule gene expression in endothelial cells and fibroblasts', *The Journal of Immunology*, 158 (7): 3285–3292.

Ehrlich P, et al. Zur Theorie der Lysenwirkung. Berlin Klin Woch. 1899;36:6.

Eicken, C., Sharma, V., Klabunde, T., Lawrenz, M.B., Hardham, J.M., Norris, S.J. and Sacchettini, J.C. (2002) 'Crystal structure of Lyme disease variable surface antigen VlsE of Borrelia burgdorferi', *Journal of Biological Chemistry*, 277(24), pp. 21691–21696.

Eikeland, R., Ljøstad, U., Mygland, A., Herlofson, K. and Løhaugen, G. (2011) 'European neuroborreliosis: Neuropsychological findings 30 months post-treatment', *European journal of neurology.*, 19 (3): 480–7.

Eikeland, R., Mygland, Å., Herlofson, K. and Ljøstad, U. (2012) 'Risk factors for a non-favorable outcome after treated European neuroborreliosis', *Acta neurologica Scandinavica.*, 127 (3): 154–60.

Eikeland, R., Mygland, Å., Herlofson, K. and Ljøstad, U. (2012) 'Risk factors for a non-favorable outcome after treated European neuroborreliosis', *Acta neurologica Scandinavica.*, 127 (3): 154–60.

Engstrom, S., Shoop, E. and Johnson, R. (1995) 'Immunoblot interpretation criteria for serodiagnosis of early Lyme disease', *Journal of clinical microbiology.*, 33 (2): 419–27.

Epand, R. (2008) 'Proteins and cholesterol-rich domains', *Biochimica et biophysica acta.*, 1778 1576–82.

Eshoo, M.W., Crowder, C.C., Rebman, A.W., Rounds, M.A., Matthews, H.E., Picuri, J.M., Soloski, M.J., Ecker, D.J., Schutzer, S.E. and Aucott, J.N. (2012) 'Direct molecular detection and Genotyping of Borrelia burgdorferi from whole blood of patients with early Lyme disease', PLoS ONE, 7(5), p. e36825.

Esteve-Gassent, M., Elliott, N. and Seshu, J. (2008) 'SodA is essential for virulence of Borrelia burgdorferi in the murine model of Lyme disease', *Molecular microbiology.*, 71 (3): 594–612.

Fahrer, H., der, van, Sauvain, M., Gern, L., Zhioua, E. and Aeschlimann, A. (1991) 'The prevalence and incidence of clinical and asymptomatic Lyme borreliosis in a population at risk', *The Journal of infectious diseases.*, 163 (2): 305–10.

Fairman J.W., Noinaj N., Buchanan S.K. (2011) 'The structural biology of beta-barrel membrane proteins: a summary of recent reports', *Curr. Opin. Struct. Biol*. 21:523–531.

Faller, M., Niederweis, M. and Schulz, G. (2004) 'The structure of a mycobacterial outer-membrane channel', *Science (New York, N.Y.).*, 303 (5661): 1189–92.

Felek, S. and Krukonis, E. (2008) 'The Yersinia pestis Ail protein mediates binding and Yop delivery to host cells required for plague virulence', *Infection and immunity.*, 77 (2): 825–36.

Ferguson, A., Chakraborty, R., Smith, B., Esser, L., der, van and Deisenhofer, J. (2002) 'Structural basis of gating by the outer membrane transporter FecA', *Science (New York, N.Y.).*, 295 (5560): 1715–9.

Ferguson, A., Hofmann, E., Coulton, J., Diederichs, K. and Welte, W. (1998) 'Siderophore-mediated iron transport: Crystal structure of FhuA with bound lipopolysaccharide', *Science (New York, N.Y.).*, 282 (5397): 2215–20.

Ferrieri, P., Burke, B. and Nelson, J. (1980) 'Production of bacteremia and meningitis in infant rats with group B streptococcal serotypes', *Infection and immunity.*, 27 (3): 1023–32.

Fikrig, E., Feng, W., Barthold, S., Telford, S. and Flavell, R. (2000) 'Arthropod- and host-specific Borrelia burgdorferi bbk32 expression and the inhibition of spirochete transmission', *Journal of immunology (Baltimore, Md. : 1950).*, 164 (10): 5344–51.

Fikrig, E., Pal, U., Chen, M., Anderson, J. and Flavell, R. (2004) 'OspB antibody prevents Borrelia burgdorferi colonization of Ixodes scapularis', *Infection and immunity.*, 72 (3): 1755–9.

Fischer, J.R., LeBlanc, K.T. and Leong, J.M. (2006) 'Fibronectin binding protein BBK32 of the Lyme disease Spirochete promotes bacterial attachment to Glycosaminoglycans', *Infection and Immunity*, 74(1), pp. 435–441.

Fischer, J., Parveen, N., Magoun, L. and Leong, J. (2003) 'Decorin-binding proteins A and B confer distinct mammalian cell type-specific attachment by Borrelia burgdorferi, the Lyme disease spirochete', *Proceedings of the National Academy of Sciences of the United States of America.*, 100 (12): 7307–12.

Fiser, A., Do, R.K. & Sali, A. (2000) 'Modeling of loops in protein structures'. Protein Science 9. 1753-1773.

Fiser, A & Sali, A. (2003) 'Modeller: generation and refinement of homology based protein structure models', Meth. Enzymol. 374: 461-91.

Fisher, M., Grimm, D., Henion, A., Elias, E., Stewart, P. and Rosa, P. (2005) 'Borrelia burgdorferi σ54 is required for mammalian infection and vector transmission but not for tick colonization', PNAS., (102): 5162-5167.

Fix, A., Peña, C. and Strickland, G. (2000) 'Racial differences in reported Lyme disease incidence', *American journal of epidemiology.*, 152 (8): 756–9.

Forst, D., Welte, W., Wacker, T. and Diederichs, K. (1998) 'Structure of the sucrose-specific porin ScrY from salmonella typhimurium and its complex with sucrose', *Nature Structural Biology*, 5 (1): 37–46.

Foulds, J. and Barrett, C. (1973) 'Characterization of Escherichia coli mutants tolerant to bacteriocin JF246: Two new classes of tolerant mutants', *Journal of bacteriology.*, 116 (2): 885–92.

Foulds, J. and Chai, T. (1978) 'Defeat of colicin tolerance in Escherichia coli ompA mutants: Evidence for interaction between colicin L-JF246 and the cytoplasmic membrane', *Journal of bacteriology.*, 133 (1): 158–64.

Franke, D. and Svergun, D.I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J. Appl. Cryst., 42, 342-346.

Fraser, C., Casjens, S., Huang, W., Sutton, G., Clayton, R., Lathigra, R., White, O., Ketchum, K., Dodson, R., Hickey, E., Gwinn, M., Dougherty, B., Tomb, J., Fleischmann, R., Richardson, D., Peterson, J., Kerlavage, A., Quackenbush, J., Salzberg, S., Hanson, M., Vugt, van, Palmer, N., Adams, M., Gocayne, J., Weidman, J., Utterback, T., Watthey, L., McDonald, L., Artiach, P., Bowman, C., Garland, S., Fuji, C., Cotton, M., Horst, K., Roberts, K., Hatch, B., Smith, H. and Venter, J. (1997) 'Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi', *Nature.*, 390 (6660): 580–6.

Freudl, R., MacIntyre, S., Degen, M. and Henning, U. (1986) 'Cell surface exposure of the outer membrane protein OmpA of Escherichia coli K-12', *Journal of molecular biology.*, 188 (3): 491–4.

Fryland, L., Wilhelmsson, P., Lindgren, P., Nyman, D., Ekerfelt, C. and Forsberg, P. (2010) 'Low risk of developing Borrelia burgdorferi infection in the south-east of Sweden after being bitten by a Borrelia burgdorferi-infected tick',

International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases., 15 (3) .

Fuchs, H., Wallich, R., Simon, M. and Kramer, M. (1994) 'The outer surface protein A of the spirochete Borrelia burgdorferi is a plasmin(ogen) receptor', *Proceedings of the National Academy of Sciences of the United States of America.*, 91 (26): 12594–8.

Fürst, B., Glatz, M., Kerl, H. and Müllegger, R. (2006) 'The impact of immunosuppression on erythema migrans. A retrospective study of clinical presentation, response to treatment and production of Borrelia antibodies in 33 patients', *Clinical and experimental dermatology.*, 31 (4): 509–14.

Galdiero, S., Capasso, D., Vitiello, M., D'Isanto, M., Pedone, C. and Galdiero, M. (2003) 'Role of surface-exposed loops of Haemophilus influenzae protein P2 in the Mitogen-Activated protein Kinase cascade', *Infection and Immunity*, 71 (5): 2798–2809.

Garcia-Monco, J., Fernandez-Villar, B. and Benach, J. (1989) 'Adherence of the Lyme disease spirochete to glial cells and cells of glial origin', *The Journal of infectious diseases.*, 160(3), pp. 497–506.

Garon, C., Dorward, D. and Corwin, M. (1989) 'Structural features of Borrelia burgdorferi--the Lyme disease spirochete: Silver staining for nucleic acids', *Scanning microscopy. Supplement.*, 3 109–15.
Gerber, M.A., Zemel, L.S. and Shapiro, E. D. (1998) 'Lyme arthritis in children: clinical epidemiology and long-term outcomes', *Pediatrics*, 102: 905-908.

Gessmann, D., Chung, Y.H., Danoff, E.J., Plummer, A.M., Sandlin, C.W., Zaccai, N.R. and Fleming, K.G. (2014) 'Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA', Proc Natl Acad Sci USA, 111(16), pp. 5878–5883.

Gil, F., Ipinza, F., Fuentes, J., Fumeron, R., Villarreal, J., Aspée, A., Mora, G., Vásquez, C. and Saavedra, C. (2007) 'The ompW (porin) gene mediates methyl viologen (paraquat) efflux in salmonella enterica serovar typhimurium', *Research in microbiology.*, 158 (6): 529–36.

Gilmore, R., Howison, R., Schmit, V., Nowalk, A., Clifton, D., Nolder, C., Hughes, J. and Carroll, J. (2007) 'Temporal expression analysis of the Borrelia burgdorferi paralogous gene family 54 genes BBA64, BBA65, and BBA66 during persistent infection in mice', *Infection and immunity.*, 75 (6): 2753–64.

Gilmore, R., Howison, R., Schmit, V. and Carroll, J. (2008) 'Borrelia burgdorferi expression of the bba64, bba65, bba66, and bba73 genes in tissues during persistent infection in mice', *Microbial pathogenesis.*, 45 355–60.

Glatz, M., Fingerle, V., Wilske, B., Ambros-Rudolph, C., Kerl, H. and Müllegger, R. (2008) 'Immunoblot analysis of the seroreactivity to recombinant Borrelia burgdorferi sensu lato antigens, including VlsE, in the long-term course of treated patients with erythema migrans', *Dermatology (Basel, Switzerland).*, 216 (2): 93–103.

Glatz, M., Golestani, M., Kerl, H. and Müllegger, R. (2006) 'Clinical relevance of different IgG and IgM serum antibody responses to Borrelia burgdorferi after antibiotic therapy for erythema migrans: Long-term follow-up study of 113 patients', *Archives of dermatology.*, 142(7), pp. 862–8.

Glatz, M., Müllegger, R., Maurer, F., Fingerle, V., Achermann, Y., Wilske, B. and Bloemberg, G. (2014) 'Detection of Candidatus Neoehrlichia mikurensis, Borrelia burgdorferi sensu lato genospecies and Anaplasma phagocytophilum in a tick population from Austria', *Ticks and tick-borne diseases.*, 5 (2): 139–44.

Glöckner, G., Lehmann, R., Romualdi, A., Pradella, S., Schulte-Spechtel, U., Schilhabel, M., Wilske, B., Sühnel, J. and Platzer, M. (2004) 'Comparative analysis of the Borrelia garinii genome', *Nucleic acids research.*, 32 (20): 6038–46.

Glöckner, G., Schulte-Spechtel, U., Schilhabel, M., Felder, M., Sühnel, J., Wilske, B. and Platzer, M. (2006) 'Comparative genome analysis: selection pressure on the Borrelia vls cassettes is essential for infectivity.', *BMC Genomics*, 7 (1): 211.

Goel, A. and Jiang, S. (2009) 'Genetic determinants of virulence, antibiogram and altered biotype among the Vibrio cholerae O1 isolates from different cholera outbreaks in India', *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases.*, 10 (6): 815–9.

Golden, M., Marra, C. and Holmes, K. (2003) 'Update on syphilis: Resurgence of an old problem', *JAMA.*, 290 (11): 1510–4.

Grab, D., Nikolskaia, O., Kim, Y., Lonsdale-Eccles, J., Ito, S., Hara, T., Fukuma, T., Nyarko, E., Kim, K., Stins, M., Delannoy, M. and Rodgers, J. (2004) 'African trypanosome interactions with an in vitro model of the human blood-brain barrier', *The Journal of parasitology.*, 90 (5): 970–9.

Grab, D. J., Perides, G., Dumler, J. S., Kim, K. J., Park, J., Kim, Y. V., Nikolskaia, O., Choi, K. S. and Stins, M. F. (2005) 'Borrelia burgdorferi, host-derived proteases, and the blood-brain barrier', *Infection and Immunity*, 73 (2): 1014–1022.

Grabski, A.C., Mehler, M., Drott, D. (2005) 'The overnight express autoinduction system: High-density cell growth and protein expression while you sleep.' Nature Methods, 2(3): 233-235.

Grassmann, A., Félix, S., Santos, dos, Amaral, M., Neto, S., Fagundes, M., Seixas, F., Silva, da, Conceição, F. and Dellagostin, O. (2012) 'Protection against lethal leptospirosis after vaccination with LipL32 coupled or coadministered with the B subunit of Escherichia coli heat-labile enterotoxin', *Clinical and vaccine immunology : CVI.*, 19 (5): 740–5.

Gray, J., Kahl, O., Robertson, J., Daniel, M., Estrada-Peña, A., Gettinby, G., Jaenson, T., Jensen, P., Jongejan, F., Korenberg, E., Kurtenbach, K. and Zeman, P. (1998) 'Lyme borreliosis habitat assessment', *Zentralblatt für Bakteriologie : international journal of medical microbiology.*, 287 (3): 211–28.

Greco, A., Gallo, A., Fusconi, C., Marinelli, G.F., Macri, M., de Vincentiis. (2012) 'Bell's palsy and autoimmunity', Autoimmun Rev., (12): 323-328.

Greco, F., Vallone, A., Apuzzo, G., Tenuta, R., Guaglianone, L. and Giraldi, C. (2003) 'Presence and indigenous nature of Lyme disease in southern Italy', *The new microbiologica.*, 26 (4): 391–4.

Greenfield, N. and Fasman, G. (1969) 'Computed circular dichroism spectra for the evaluation of protein conformation', *Biochemistry.*, 8(10), pp. 4108–4116.

Gregory, E.M., Yost, F.J. and Fridovich, I. (1973) 'Superoxide dismutase of Escherichia coli: intracellular localization and functions', *J Bacteriol,* 115 (3): 987-981.

Grimm, D., Tilly, K., Byram, R., Stewart, P., Krum, J., Bueschel, D., Schwan, T., Policastro, P., Elias, A. and Rosa, P. (2004) 'Outer-surface protein C of the Lyme disease spirochete: A protein induced in ticks for infection of mammals', *Proceedings of the National Academy of Sciences of the United States of America.*, 101 (9): 3142–7.

Guinier A (1939) La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques. Ann Phys 12: 161–237.

Guo, X., Booth, C. J., Paley, M. A., Wang, X., DePonte, K., Fikrig, E., Narasimhan, S. and Montgomery, R. R. (2009) 'Inhibition of Neutrophil function by Two tick salivary proteins', *Infection and Immunity*, 77 (6): 2320–2329.

Guttenplan, S. and Kearns, D. (2013) 'Regulation of flagellar motility during biofilm formation', FEMS microbiology reviews., 37(6), pp. 849–71.

Gwakisa, P., Yoshihara, K., Long, T., Gotoh, H., Amano, F. and Momotani, E. (2001) 'Salivary gland extract of Rhipicephalus appendiculatus ticks inhibits in vitro transcription and secretion of cytokines and production of nitric oxide by LPS-stimulated JA-4 cells', *Veterinary parasitology.*, 99 (1): 53–61.

Haake, D. (2000) 'Spirochaetal lipoproteins and pathogenesis', *Microbiology (Reading, England).*, 146 1491–504.

Hallström, T., Haupt, K., Kraiczy, P., Hortschansky, P., Wallich, R., Skerka, C. and Zipfel, P. (2010) 'Complement regulator-acquiring surface protein 1 of Borrelia burgdorferi binds to human bone morphogenic protein 2, several extracellular matrix proteins, and plasminogen', *The Journal of infectious diseases.*, 202 (3): 490–8.

Halperin, J. (2011) 'Nervous system Lyme diease: Is there a controversy?' Semin Neurol., 31(3): 317-324.

Hansen, K. and Lebech, A. (1992) 'The clinical and epidemiological profile of Lyme neuroborreliosis in Denmark 1985-1990. A prospective study of 187 patients with Borrelia burgdorferi specific intrathecal antibody production', *Brain : a journal of neurology.*, 115 399–423.

Harris, G., Wenjiang, M., Maurer, L.M., Potts, J.R., Mosher, D.F. (2014) 'Borrelia burgdorferi protein BBK32 binds to soluble Fibronectin via the n-terminal 70-kDa region, causing Fibronectin to undergo conformational extension', *Journal of Biological Chemistry*, 289(32), pp. 22490–22499.

Hassan, H.M. and Fridovich, I. (1979) 'Intracellular production of superoxide radical and of hydrogen peroxide by redox active compounds', *Arch Biochem Biophys*. 196(2): 385-395.

Haupt, K., Kraiczy, P., Wallich, R., Brade, V., Skerka, C. and Zipfel, P. (2007) 'Binding of human factor h-related protein 1 to serum-resistant Borrelia burgdorferi is mediated by borrelial complement regulator-acquiring surface proteins', *The Journal of infectious diseases.*, 196 (1): 124–33.

Hayat, S. and Elofsson, A. (2012) 'Ranking models of transmembrane β-barrel proteins using Z-coordinate predictions', *Bioinformatics (Oxford, England).*, 28 (12): 90–96.

Health Protection Agency (2013) 'Lyme borreliosis epidemiology and surveillance', available at https://www.gov.uk/government/publications/lyme-borreliosis-epidemiology/lyme-borreliosis-epidemiology-and-surveillance, accessed 29 March 2016.

Heffernan, E.J., Harwood, J., Fierer, J. and Guiney, D. (1992) 'The salmonella typhimurium virulence plasmid complement resistance gene rck is homologous to a family of virulence-related outer membrane protein genes, including pagC and ail', *Journal of Bacteriology*, 174(1), pp. 84–91.

Hellerström, S. (1930) 'Erythema chronicum migrans Afzelii', Acta Dermato-Benereologica., (11): 315-321.

Hellwage, J., Meri, T., Heikkilä, T., Alitalo, A., Panelius, J., Lahdenne, P. and Seppälä, I. (2000) 'The complement regulator factor H binds to the surface protein OspE of Borrelia burgdorferi', *The Journal of biological chemistry.*, 276(11), pp. 8427–35.

Hengge, U., Tannapfel, A., Tyring, S., Erbel, R., Arendt, G. and Ruzicka, T. (2003) 'Lyme borreliosis', *The Lancet. Infectious diseases.*, 3 (8): 489–500.

Henikoff, S.; Henikoff, J.G. (1992). 'Amino Acid Substitution Matrices from Protein Blocks'. *PNAS*. **89** (22): 10915–10919.

Henningsson, A., Malmvall, B., Ernerudh, J., Matussek, A. and Forsberg, P. (2010) 'Neuroborreliosis--an epidemiological, clinical and healthcare cost study from an endemic area in the south-east of Sweden', *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases.*, 16 (8): 1245–51.

Herzberger, P., Siegel, C., Skerka, C., Fingerle, V., Schulte-Spechtel, U., Dam, van, Wilske, B., Brade, V., Zipfel, P., Wallich, R. and Kraiczy, P. (2007) 'Human pathogenic Borrelia spielmanii sp. Nov. Resists complement-mediated killing by direct binding of immune regulators factor H and factor h-like protein 1', *Infection and immunity.*, 75 (10): 4817–25.

Hinnebusch, B., Jarrett, C., Callison, J., Gardner, D., Buchanan, S. and Plano, G. (2011) 'Role of the Yersinia pestis Ail protein in preventing a protective polymorphonuclear leukocyte response during bubonic plague', *Infection and immunity.*, 79 (12): 4984–9.

Hofmann, K. and Stoffel, W. (1993) 'TMbase - A database of membrane spanning proteins segments', *Biol. Chem*, 374(166).

Hong, H., Patel, D., Tamm, L. and den, van (2006) 'The outer membrane protein OmpW forms an eight-stranded beta-barrel with a hydrophobic channel', *The Journal of biological chemistry.*, 281 (11): 7568–77.
Hovind-Hougen, K. (1984) 'Ultrastructure of spirochetes isolated from Ixodes ricinus and Ixodes dammini', 57 (4): 543–548.

Hovius, J., Schuijt, T., Groot, de, Roelofs, J., Oei, G., Marquart, J., Beer, de, van, 't, der, van, Ramamoorthi, N., Fikrig, E. and Dam, van (2008) 'Preferential protection of Borrelia burgdorferi sensu stricto by a Salp15 homologue in Ixodes ricinus saliva', *The Journal of infectious diseases.*, 198 (8): 1189–97.

Hsieh, Y.-F., Liu, H.-W., Hsu, T.-C., Wei, J. C.-C., Shih, C.-M., Krause, P. J. and Tsay, G. J. (2007) 'Serum reactivity against Borrelia burgdorferi OspA in patients with rheumatoid Arthritis▽', 14 (11): 1437–1441.
Huang, X., Nakagawa, T., Tamura, A., Link, K. and Koide, A. (2001) 'Formation of the single-layer beta-sheet of Borrelia burgdorferi OspA in the absence of the c-terminal capping globular domain', *Journal of molecular biology.*, 308 (2): 367–75.

Hu, L., Pratt, S., Perides, G., Katz, L., Rogers, R. and Klempner, M. (1997) 'Isolation, cloning, and expression of a 70-kilodalton plasminogen binding protein of Borrelia burgdorferi', *Infection and immunity.*, 65(12), pp. 4989–95.

Huang, X., Yang, X., Luft, B. and Koide, S. (1998) 'NMR identification of epitopes of Lyme disease antigen OspA to monoclonal antibodies', *Journal of molecular biology.*, 281 (1): 61–7.

Hubálek, Z. (2009) 'Epidemiology of lyme borreliosis', *Current problems in dermatology.*, 37 31–50.

Huegli, D., Moret, J., Rais, O., Moosmann, Y., Erard, P., Malinverni, R. and Gern, L. (2011) 'Prospective study on the incidence of infection by Borrelia burgdorferi sensu lato after a tick bite in a highly endemic area of Switzerland', *Ticks and tick-borne diseases.*, 2 (3): 129–36.

Hughes, J., Nolder, C., Nowalk, A., Clifton, D., Howison, R., Schmit, V., Gilmore, R. and Carroll, J. (2008) 'Borrelia burgdorferi surface-localized proteins expressed during persistent murine infection are conserved among diverse Borrelia spp', *Infection and immunity.*, 76 (6): 2498–511.

Hugli TE. (1986) 'Biochemistry and biology of anaphylatoxins'. *Complement*. 3:111–127.

Huppertz, H., Böhme, M., Standaert, S., Karch, H. and Plotkin, S. (1999) 'Incidence of Lyme borreliosis in the Würzburg region of Germany', *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology.*, 18 (10): 697–703.

Huttner, W. and Zimmerberg, J. (2001) 'Implications of lipid microdomains for membrane curvature, budding and fission', *Current opinion in cell biology.*, 13 (4): 478–84.

Hwang, P., Choy, W., Lo, E., Chen, L., Forman-Kay, J., Raetz, C., Privé, G., Bishop, R. and Kay, L. (2002) 'Solution structure and dynamics of the outer membrane enzyme PagP by NMR', *Proceedings of the National Academy of Sciences of the United States of America.*, 99 (21): 13560–5.

Indest, K., Howell, J., Jacobs, M., Scholl-Meeker, D., Norris, S. and Philipp, M. (2001) 'Analysis of Borrelia burgdorferi vlsE gene expression and recombination in the tick vector', *Infection and immunity.*, 69 (11): 7083–90.

Inoue, T., Yoshikawa, K., Tada, M., Nakamura, T., Hinoshita, F. (2010) 'Two cases of weil's disease with acute renal failure in the central Tokyo metropolitan area', *Clinical Nephrol*. 73(1):76-80.

Inouye, S., Wang, S., Sekizawa, J. and Halegoua, S. (1977) 'Amino acid sequence for the peptide extension on the prolipoprotein of the Escherichia coli outer membrane', *Proceedings of the National Academy of Sciences of the United States of America.*, 74(3), pp. 1004–8.

Iranzo, O. (2011) 'Manganese complexes displaying superoxide dismutase activity: A balance between different factors', *Bioorganic chemistry.*, 39 (2): 73–87.

Jackson, T. and Brunold, T. (2004) 'Combined spectroscopic/computational studies on Fe- and Mn-dependent superoxide dismutases: Insights into second-sphere tuning of active site properties', *Accounts of chemical research.*, 37 (7): 461–70.

Jacques, D.A. and Trewhella, J. (2010) 'Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls', *Protein Science*, 19(4), pp. 642–657.

Jang, M., Kim, J., Chung, Y., Lee, W., Shin, Y., Lee, J., Kim, D. and Park, Y. (2011) 'Dendritic cells stimulated with outer membrane protein A (OmpA) of salmonella typhimurium generate effective anti-tumor immunity', *Vaccine.*, 29 (13): 2400–10.

Jiao, Y., Wen, B., Chen, M., Niu, D., Zhang, J. and Qiu, L. (2006) 'Analysis of immunoprotectivity of the recombinant OmpA of Rickettsia heilongjiangensis', *Annals of the New York Academy of Sciences.*, 1063: 261–5.

Johnson, E. and Wessling-Resnick, M. (2011) 'Iron metabolism and the innate immune response to infection', *Microbes and infection / Institut Pasteur.*, 14 (3): 207–16.

Jokiranta, S.T., Zipfel, P.F., Fremeaux-Bacchi, V., Taylor, M.C., Goodship, T.J.H. and Noris, M. (2016) 'Where next with atypical hemolytic uremic syndrome?', *Molecular Immunology*, 44(16), pp. 3889–3900.

Jolley, K. A., Appleby, L., Wright, J. C., Christodoulides, M., & Heckels, J. E. (2001). Immunization with Recombinant Opc Outer Membrane Protein from Neisseria meningitidis: Influence of Sequence Variation and Levels of Expression on the Bactericidal Immune Response against Meningococci. Infection and Immunity, 69(6), 3809–3816.

Jones, K., Seward, R., Ben-Menachem, G., Glickstein, L., Costello, C. and Steere, A. (2009) 'Strong IgG antibody responses to Borrelia burgdorferi glycolipids in patients with Lyme arthritis, a late manifestation of the infection', *Clinical immunology (Orlando, Fla.).*, 132 (1): 93–102.

Kajiya, M., Komatsuzawa, H., Papantonakis, A., Seki, M., Makihira, S., Ouhara, K., Kusumoto, Y., Murakami, S., Taubman, M. and Kawai, T. (2011) 'Aggregatibacter actinomycetemcomitans Omp29 is associated with bacterial entry to gingival epithelial cells by F-actin rearrangement', *PloS one.*, 6 (4): e18287.

Kalish, R., McHugh, G., Granquist, J., Shea, B., Ruthazer, R. and Steere, A. (2001) 'Persistence of immunoglobulin M or immunoglobulin G antibody responses to Borrelia burgdorferi 10-20 years after active Lyme disease', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.*, 33 (6): 780–5.

Källberg, M., Wang, H., Peng, J., Lu, H. and Xu, J. (2012) 'Template-based protein structure modeling using the RaptorX web server', *Nature protocols.*, 7(8), pp. 1511–22.

Kang, Y., He, Y., Zhao, M. and Li, W. (2011) 'Structures of native and Fe-substituted SOD2 from Saccharomyces cerevisiae', *Acta crystallographica. Section F, Structural biology and crystallization communications.*, 67 1173–8.

Keele B. B., Jr., McCord J. M., Fridovich I. (1970) Superoxide dismutase from *Escherichia coli* B. A new manganese-containing enzyme. J. Biol. Chem. 245, 6176–6181

Keen, J.N., Caceres, I., Eliopoulos, E.E., Zagalsky, P.F. and Findlay, J.B.C. (1991) 'Complete sequence and model for the C1 subunit of the carotenoprotein crustacyanin, and model for the dimer, beta-crustacyanin, formed from the C1 and A2 subunits with astaxanthin', *European Journal of Biochemistry*, 202(1), pp. 31–40.

Kelley, L.A. (2015) 'The Phyre2 web portal for protein modeling, prediction and analysis', *Nature Protocols*, 10(6), pp. 845–858

Kelly, S., Jess, T. and Price, N. (2005) 'How to study proteins by circular dichroism', *Biochimica et biophysica acta.*, 1751(2), pp. 119–39.

Kemper C, Atkinson JP, Hourcade DE. (2010) 'Properdin: emerging roles of a pattern-recognition molecule'. *Annu Rev Immunol.* 28:131–155

Kenedy, M., Luthra, A., Anand, A., Dunn, J. and Radolf, J. (2013) 'Structural modeling and physicochemical characterization provide evidence that P66 forms a β-barrel in the Borrelia burgdorferi outer membrane', *Journal of bacteriology.*, 196 (4): 859–72.

Kenedy, M. and Lenhart, T. (2012) 'The role of Borrelia burgdorferi outer surface proteins', *FEMS immunology and medical microbiology.*, 66 (1): 1–19.

Khalid, S., Bond, P.J., Deol, S.S. and Sansom, M.S.P. (2006) 'Modeling and simulations of a bacterial outer membrane protein: OprF from Pseudomonas aeruginosa', *Proteins: Structure, Function, and Bioinformatics*, 63(1), pp. 6–15.

Khatchikian, C., Nadelman, R., Nowakowski, J., Schwartz, I., Wormser, G. and Brisson, D. (2014) 'Evidence for strain-specific immunity in patients treated for early lyme disease', *Infection and immunity.*, 82 (4): 1408–13.

Kim, K. (1987) 'Effect of antimicrobial therapy for experimental infections due to group B Streptococcus on mortality and clearance of bacteria', *The Journal of infectious diseases.*, 155 (6): 1233–41.

Kim, K. (2001) 'Escherichia coli translocation at the blood-brain barrier', *Infection and immunity.*, 69(9), pp. 5217–22.

Kim, K.S. (2006) 'Microbial translocation of the blood–brain barrier', *International Journal for Parasitology*, 36(5), pp. 607–614

Kim, K., Kim, K., Choi, J., Lim, J., Lee, J., Hwang, S. and Ryu, S. (2010) 'Outer membrane proteins A (OmpA) and X (OmpX) are essential for basolateral invasion of Cronobacter sakazakii', *Applied and environmental microbiology.*, 76 (15): 5188–98.

Kim, K. S., Itabashi, H., Gemski, P., Sadoff, J., Warren, R. L. and Cross, A. S. (1992) 'The K1 capsule is the critical determinant in the development of Escherichia coli meningitis in the rat', *J Clin Invest.*, 90 (3): 897–905.

Kindstrand, E., Nilsson, B., Hovmark, A., Pirskanen, R. and Asbrink, E. (1997) 'Peripheral neuropathy in acrodermatitis chronica atrophicans - a late Borrelia manifestation', *Acta neurologica Scandinavica.*, 95 (6): 338–45.

Klempner, M., Hu, L., Evans, J., Schmid, C., Johnson, G., Trevino, R., Norton, D., Levy, L., Wall, D., McCall, J., Kosinski, M. and Weinstein, A. (2001) 'Two controlled trials of antibiotic treatment in patients with persistent symptoms and a history of Lyme disease', *The New England journal of medicine.*, 345 (2): 85–92.

Kobryn, K. and Chaconas, G. (2002) 'ResT, a telomere resolvase encoded by the Lyme disease spirochete', *Molecular cell.*, 9 (1): 195–201.

Koch, M., Vachette, P. and Svergun, D. (2003) 'Small-angle scattering: A view on the properties, structures and structural changes of biological macromolecules in solution', *Quarterly reviews of biophysics.*, 36(2), pp. 147–227.

Koebnik, R., Locher, K. and Gelder, V. (2000) 'Structure and function of bacterial outer membrane proteins: Barrels in a nutshell', *Molecular microbiology.*, 37 (2): 239–53.

Kolodziejek, A., Schnider, D., Rohde, H., Wojtowicz, A., Bohach, G., Minnich, S. and Hovde, C. (2010) 'Outer membrane protein X (Ail) contributes to Yersinia pestis virulence in pneumonic plague and its activity is dependent on the lipopolysaccharide core length', *Infection and immunity.*, 78 (12): 5233–43.

Kolodziejek, A., Sinclair, D., Seo, K., Schnider, D., Deobald, C., Rohde, H., Viall, A., Minnich, S., Hovde, C. and Bohach, G. (2007) 'Phenotypic characterization of OmpX, an Ail homologue of Yersinia pestis KIM', *Microbiology (Reading, England).*, 153 2941–51.

Konarev, P. & Volkov, V., Sokolova, A., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS - a Windows-PC based system for small-angle scattering data analysis. J Appl Cryst. 36, 1277-1282.

Korenberg, E., Vorobyeva, N., Moskvitina, H. and LYa, G. (1996) 'Prevention of borreliosis in persons bitten by infected ticks', *Infection.*, 24 (2): 187–9.

Koronakis, V., Sharff, A., Luisi, B. and Hughes, C. (2000) 'Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export', *Nature.*, 405 (6789): 914–9.

Kort, de, Bolton, A., Martin, G., Stephen, J. and van, de (1994) 'Invasion of rabbit ileal tissue by Enterobacter cloacae varies with the concentration of OmpX in the outer membrane', *Infection and immunity.*, 62 (11): 4722–6.

Kovacs-Simon, A., Titball, R.W. and Michell, S.L. (2011) 'Lipoproteins of bacterial Pathogens', *Infection and Immunity*, 79(2), pp. 548–561.

Kowalski, T., Tata, S., Berth, W., Mathiason, M. and Agger, W. (2010) 'Antibiotic treatment duration and long-term outcomes of patients with early lyme disease from a lyme disease-hyperendemic area', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.*, 50 (4): 512–20.

Kozin, M. & Svergun, D. (2001) Automated matching of high- and low-resolution structural models *J Appl Cryst.* 34, 33-41.

Kraiczy, P., Hartmann, K., Hellwage, J., Skerka, C., Kirschfink, M., Brade, V., Zipfel, P., Wallich, R. and Stevenson, B. (2004) 'Immunological characterization of the complement regulator factor h-binding CRASP and Erp proteins of Borrelia burgdorferi', *International journal of medical microbiology : IJMM.*, 293, pp. 152–7.

Kraiczy, P., Skerka, C., Kirschfink, M., Zipfel, P. and Brade, V. (2001) 'Mechanism of complement resistance of pathogenic Borrelia burgdorferi isolates', *International immunopharmacology.*, 1 (3): 393–401.

Kreusch, A., Neubüser, A., Schiltz, E., Weckesser, J. and Schulz, G. (1994) 'Structure of the membrane channel porin from Rhodopseudomonas blastica at 2.0 A resolution', *Protein science : a publication of the Protein Society.*, 3 (1): 58–63.

Krishnan, S. and Prasadarao, N. (2012) 'Outer membrane protein A and OprF: Versatile roles in gram-negative bacterial infections', *The FEBS journal.*, 279 (6): 919–31.

Kristoferitsch, W., Sluga, E., Graf, M., Partsch, H., Neumann, R., Stanek, G. and Budka, H. (1988) 'Neuropathy associated with acrodermatitis chronica atrophicans. Clinical and morphological features', *Annals of the New York Academy of Sciences.*, 539 35–45.

Krupka, M., Raska, M., Belakova, J., Horynova, M., Novotny, R. and Weigl, E. (2008) 'Biological aspects of Lyme disease spirochetes: Unique bacteria of the Borrelia burgdorferi species group', *Biomedical papers of the Medical Faculty of the University Palacký, Olomouc, Czechoslovakia.*, 151 (2): 175–86.

Krupka, M., Zachova, K., Weigl, E. and Raska, M. (2011) 'Prevention of lyme disease: Promising research or sisyphean task?', *Archivum immunologiae et therapiae experimentalis.*, 59 (4): 261–75.

Krupp, L., Hyman, L., Grimson, R., Coyle, P., Melville, P., Ahnn, S., Dattwyler, R. and Chandler, B. (2003) 'Study and treatment of post Lyme disease (STOP-LD): A randomized double masked clinical trial', *Neurology.*, 60 (12): 1923–30.

Kubes, M., Kocáková, P., Slovák, M., Sláviková, M., Fuchsberger, N. and Nuttall, P. (2002) 'Heterogeneity in the effect of different ixodid tick species on human natural killer cell activity', *Parasite immunology.*, 24 (1): 23–8.

Kudryashev, M., Cyrklaff, M., Baumeister, W., Simon, M., Wallich, R. and Frischknecht, F. (2009) 'Comparative cryo-electron tomography of pathogenic Lyme disease spirochetes', *Molecular microbiology.*, 71 (6): 1415–34.

Kuehn, B. (2013) 'CDC estimates 300, 000 US cases of Lyme disease annually', *JAMA.*, 310 (11): 1110.

Kumar, M., Kaur, S., Kariu, T., Yang, X., Bossis, I., Anderson, J. and Pal, U. (2011) 'Borrelia burgdorferi BBA52 is a potential target for transmission blocking Lyme disease vaccine', *Vaccine.*, 29 (48): 9012–9.

Kumaran, D., Eswaramoorthy, S., Luft, B.J., Koide, S., Dunn, J.J., Lawson, C.L. and Swaminathan, S. (2001) 'Crystal structure of outer surface protein C (OspC) from the Lyme disease spirochete, Borrelia burgdorferi', *EMBO*, 20(5), pp. 971–978.

Kung, F., Anguita, J. and Pal, U. (2013) 'Borrelia burgdorferi and tick proteins supporting pathogen persistence in the vector', *Future Microbiology*, 8, pp. 41–56.

Labandeira-Rey, M., Seshu, J. and Skare, J. (2003) 'The absence of linear plasmid 25 or 28-1 of Borrelia burgdorferi dramatically alters the kinetics of experimental infection via distinct mechanisms', *Infection and immunity.*, 71(8), pp. 4608–13.

Labandeira-Rey, M. and Skare, J. (2000) 'Decreased infectivity in Borrelia burgdorferi strain B31 is associated with loss of linear plasmid 25 or 28-1', *Infection and immunity.*, 69 (1): 446–55.

Laemmli, U.K. (1970) 'Cleavage of structural proteins during the assembly of the head of Bacteriophage T4', Nature, 227(5259), pp. 680–685.

LaFrentz, B., Shoemaker, C. and Klesius, P. (2011) 'Immunoproteomic analysis of the antibody response obtained in Nile tilapia following vaccination with a Streptococcus iniae vaccine', *Veterinary microbiology.*, 152, pp. 346–52.

Lagal, V., Portnoï, D., Faure, G., Postic, D. and Baranton, G. (2006) 'Borrelia burgdorferi sensu stricto invasiveness is correlated with OspC-plasminogen affinity', *Microbes and infection / Institut Pasteur.*, 8 (3): 645–52.

Lah, M., Dixon, M., Pattridge, K., Stallings, W., Fee, J. and Ludwig, M. (1995) 'Structure-function in Escherichia coli iron superoxide dismutase: Comparisons with the manganese enzyme from Thermus thermophilus', *Biochemistry.*, 34(5), pp. 1646–60.

Lakos, A. and Solymosi, N. (2009) 'Maternal Lyme borreliosis and pregnancy outcome', *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases.*, 14 (6).

Lam, T., Nguyen, T., Montgomery, R., Kantor, F., Fikrig, E. and Flavell, R. (1994) 'Outer surface proteins E and F of Borrelia burgdorferi, the agent of Lyme disease', *Infection and immunity.*, 62(1), pp. 290–8.

Lambris, J., Ricklin, D. and Geisbrecht, B. (2008) 'Complement evasion by human pathogens', *Nature reviews. Microbiology.*, 6 (2): 132–42.

Landau, E. and Rosenbusch, J. (1996) 'Lipidic cubic phases: A novel concept for the crystallization of membrane proteins', *Proceedings of the National Academy of Sciences of the United States of America.*, 93(25), pp. 14532–5.

Lawson, C., Yung, B., Barbour, A. and Zückert, W. (2006) 'Crystal structure of neurotropism-associated variable surface protein 1 (Vsp1) of Borrelia turicatae', *Journal of bacteriology.*, 188(12), pp. 4522–30.

Lebech, A. (2002) 'Polymerase chain reaction in diagnosis of Borrelia burgdorferi infections and studies on taxonomic classification', *APMIS. Supplementum.*, 105 1–40.

Lee, S., Lee, J., Park, H., Jang, W., Koh, S., Yang, Y., Kim, B., Kook, Y. and Park, K. (2003) 'Differentiation of Borrelia burgdorferi sensu lato through groEL gene analysis', *FEMS microbiology letters.*, 222 (1): 51–7.

Lees, J.G., Miles, A.J., Wien, F., and Wallace, B.A. (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics* **22** 1955-1962.

Lenhart, T. and Akins, D. (2009) 'Borrelia burgdorferi locus BB0795 encodes a BamA orthologue required for growth and efficient localization of outer membrane proteins', *Molecular microbiology.*, 75 (3): 692–709.

Leong, J., Wang, H., Magoun, L., Field, J., Morrissey, P., Robbins, D., Tatro, J., Coburn, J. and Parveen, N. (1998) 'Different classes of proteoglycans contribute to the attachment of Borrelia burgdorferi to cultured endothelial and brain cells', *Infection and immunity.*, 66 (3): 994–9.

Lesnyak, O., Laikovskaya, E., Kufko, I., Bruinink, H., Baranova, N. and Rijpkema, S. (1998) 'Clinical features of Lyme borreliosis in the middle Urals and distribution of Borrelia burgdorferi sensu lato species in local Ixodes persulcatus ticks', *Zentralblatt für Bakteriologie : international journal of medical microbiology.*, 288 (1): 111–9.

Li, C., Xu, H., Zhang, K. and Liang, F. (2010) 'Inactivation of a putative flagellar motor switch protein FliG1 prevents Borrelia burgdorferi from swimming in highly viscous media and blocks its infectivity', *Molecular microbiology.*, 75 (6): 1563–76.

Li, W., Wang, H., Lei, C., Ying, T. and Tan, X. (2015) 'Manganese superoxide dismutase from human pathogen Clostridium difficile', Amino acids., 47(5), pp. 987–95.

Li, X., Neelakanta, G., Liu, X., Beck, D., Kantor, F., Fish, D., Anderson, J. and Fikrig, E. (2007) 'Role of outer surface protein D in the Borrelia burgdorferi life cycle', *Infection and immunity.*, 75 (9): 4237–44.

Li, X., Pal, U., Ramamoorthi, N., Liu, X., Desrosiers, D., Eggers, C., Anderson, J., Radolf, J. and Fikrig, E. (2006) 'The Lyme disease agent Borrelia burgdorferi requires BB0690, a Dps homologue, to persist within ticks', *Molecular microbiology.*, 63 (3): 694–710.

Liang, F., Aberer, E., Cinco, M., Gern, L., Hu, C., Lobet, Y., Ruscio, M., Voet, P., Weynants, V. and Philipp, M. (2000) 'Antigenic conservation of an immunodominant invariable region of the VlsE lipoprotein among European pathogenic genospecies of Borrelia burgdorferi SL', *The Journal of infectious diseases.*, 182(5), pp. 1455–62.

Liang, F., Nelson, F. and Fikrig, E. (2002) 'DNA microarray assessment of putative Borrelia burgdorferi lipoprotein genes', *Infection and immunity.*, 70 (6): 3300–3.

Liang, F., Xu, Q., Sikdar, R., Xiao, Y., Cox, J. and Doerrler, W. (2010) 'BB0250 of Borrelia burgdorferi is a conserved and essential inner membrane protein required for cell division', *Journal of bacteriology.*, 192 (23): 6105–15.

Liao, J., Li, H., Zeng, W., Sauer, D., Belmares, R. and Jiang, Y. (2012) 'Structural insight into the ion-exchange mechanism of the sodium/calcium exchanger', Science (New York, N.Y.)., 335(6069), pp. 686–90.

Lipsker, D., Hansmann, Y., Limbach, F., Clerc, C., Tranchant, C., Grunenberger, F., Caro-Sampara, F., Attali, P., Frey, M., Kubina, M., Piémont, Y., Sibilia, J., Jaulhac, B. and Study, G. (2001) 'Disease expression of Lyme borreliosis in northeastern France', *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology.*, 20 (4): 225–30.

Liu, P., Ewis, H., Huang, Y., Lu, C., Tai, P. and Weber, I. (2007) 'Structure of Bacillus subtilis superoxide dismutase', Acta Crystallogr Sect F Struct Biol Cryst Commun, 63(Pt 12), pp. 1003–1007.

Livengood, J. and Gilmore, R. (2006) 'Invasion of human neuronal and glial cells by an infectious strain of Borrelia burgdorferi', *Microbes and infection / Institut Pasteur.*, 8 2832–40.
Ljøstad, U. and Mygland, A. (2009) 'Remaining complaints 1 year after treatment for acute Lyme neuroborreliosis; frequency, pattern and risk factors', *European journal of neurology.*, 17 (1): 118–23.

Liveris, D., Wang, G., Girao, G., Byrne, D., Nowakowski, J., McKenna, D., Nadelman, R., Wormser, G. and Schwartz, I. (2002) 'Quantitative detection of Borrelia burgdorferi in 2-millimeter skin samples of erythema migrans lesions: Correlation of results with clinical and laboratory findings', *Journal of clinical microbiology.*, 40(4), pp. 1249–53.

Lledó, L., Gegúndez, M., Giménez-Pardo, C., Álamo, R., Fernández-Soto, P., Nuncio, M. and Saz, J. (2014) 'A seventeen-year epidemiological surveillance study of Borrelia burgdorferi infections in two provinces of northern Spain', *International journal of environmental research and public health.*, 11 (2): 1661–72.

Llobet, E., Catalina, M., Giménez, P. and Bengoechea, J.A. (2009) 'Klebsiella pneumoniae OmpA confers resistance to Antimicrobial peptides', *Antimicrobial Agents and Chemotherapy*, 53 (1): 298–302.

Lobley, A., Whitmore, L. and Wallace, B. (2002) 'DICHROWEB: An interactive website for the analysis of protein secondary structure from circular dichroism spectra', *Bioinformatics*, 18(1), pp. 211–2.

Love, D., Redwin, J. and Norris, J. (2002) 'Cloning and expression of the superoxide dismutase gene of the feline strain of Porphyromonas gingivalis: Immunological recognition of the protein by cats with periodontal disease', *Veterinary microbiology.*, 86(3), pp. 245–56.

Ludwig, M., Metzger, A., Pattridge, K. and Stallings, W. (1991) 'Manganese superoxide dismutase from Thermus thermophilus. A structural model refined at 1.8 A resolution', Journal of molecular biology., 219(2), pp. 335–58.

Lugtenberg, B. and Alphen, V. (1983) 'Molecular architecture and functioning of the outer membrane of Escherichia coli and other gram-negative bacteria', *Biochimica et biophysica acta.*, 737 (1): 51–115.

Madico, G., Welsch, J., Lewis, L., McNaughton, A., Perlman, D., Costello, C., Ngampasutadol, J., Vogel, U., Granoff, D. and Ram, S. (2006) 'The meningococcal vaccine candidate GNA1870 binds the complement regulatory protein factor H and enhances serum resistance', *Journal of immunology (Baltimore, Md. : 1950).*, 177(1), pp. 501–10.

Magoun, L., Zückert, W., Robbins, D., Parveen, N., Alugupalli, K., Schwan, T., Barbour, A. and Leong, J. (2000) 'Variable small protein (Vsp)-dependent and Vsp-independent pathways for glycosaminoglycan recognition by relapsing fever spirochaetes', *Molecular microbiology.*, 36(4), pp. 886–97.

Maiwald, M., Oehme, R., March, O., Petney, T., Kimmig, P., Naser, K., Zappe, H., Hassler, D. and Knebel, von (1998) 'Transmission risk of Borrelia burgdorferi sensu lato from Ixodes ricinus ticks to humans in southwest Germany', *Epidemiology and infection.*, 121 (1): 103–8.

Makabe, K., Tereshko, V., Gawlak, G., Yan, S. and Koide, S. (2006) 'Atomic-resolution crystal structure ofBorrelia burgdorferiouter surface protein A via surface engineering', Protein Science, 15(8), pp. 1907–1914.

Manavalan, P. and Johnson, C.W. (1987) 'Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra', *Analytical Biochemistry*, 167(1), pp. 76–85.

Mandel, M. and Higa, A. (1970) 'Calcium-dependent bacteriophage DNA infection', *Journal of molecular biology.*, 53(1), pp. 159–62.

Mandelli, F., Cairo, F., Citadini, A., Büchli, F., Alvarez, T., Oliveira, R., Leite, V., Leme, P., Mercadante, A. and Squina, F. (2013) 'The characterization of a thermostable and cambialistic superoxide dismutase from Thermus filiformis', *Letters in applied microbiology.*, 57(1), pp. 40–6.

Maraspin, V., Cimperman, J., Lotric-Furlan, S., Ruzić-Sabljić, E., Jurca, T., Picken, R. and Strle, F. (2002) 'Solitary borrelial lymphocytoma in adult patients', *Wiener klinische Wochenschrift.*, 114 515–23.

Margos, G., Vollmer, S., Ogden, N. and Fish, D. (2011) 'Population genetics, taxonomy, phylogeny and evolution of Borrelia burgdorferi sensu lato', *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases.*, 11 (7): 1545–63.

Marra, A. and Brigham, D. (2001) 'Streptococcus pneumoniae causes experimental Meningitis following intranasal and Otitis media infections via a Nonhematogenous route', *The journal of experimental medicine*, 69 (12): 1173–1183.

Marti-Renom, M.A., Stuart, A., Fiser, A., Sánchez, R., Melo, F., Sali, A. (2000) 'Comparative protein structure modeling of genes and genomes'. Annu. Rev. Biophys. Biomol. Struct. 29, 291-325.

Martin, D., Brodeur, B., Hamel, J., Couture, F., Alwis, de, Lian, Z., Andrews, D. and Ellis, R. (2000) 'Candidate Neisseria meningitidis NspA vaccine', *Journal of biotechnology.*, 83: 27–31.

Martin, D., Cadieux, N., Hamel, J. and Brodeur, B. (1997) 'Highly conserved Neisseria meningitidis surface protein confers protection against experimental infection', *The Journal of experimental medicine.*, 185 (7): 1173–83.

Martin M. E., Byers B. R., Olson M. O., Salin M. L., Arceneaux J. E., Tolbert C. (1986) A *Streptococcus mutans* superoxide dismutase that is active with either manganese or iron as a cofactor. J. Biol. Chem. 261, 9361–9367.

Masocha, W., Rottenberg, M. and Kristensson, K. (2007) 'Migration of African trypanosomes across the blood-brain barrier', *Physiology & behavior.*, 92 110–4.

Mathiopoulos, C., Mueller, J., Slack, F., Murphy, C., Patankar, S., Bukusoglu, G. and Sonenshein, A. (1991) 'A Bacillus subtilis dipeptide transport system expressed early during sporulation', *Molecular microbiology.*, 5 (8): 1903–13.

Maurya, S. and Mahalakshmi, R. (2013) 'Modulation of human mitochondrial voltage-dependent anion channel 2 (hVDAC-2) structural stability by cysteine-assisted barrel-lipid interactions', *The Journal of biological chemistry.*, 288 (35): 25584–92.

Mawanda, F and Wallace, R. (2013) 'Can infections cause Alzheimer's disease?' Epidemiol Rev., 35(1): 161-180.

M'bomeyo, L., Hedelin, G. and Lipsker, D. (2003) '[The level of knowledge of general practitioners regarding the early phase of Lyme borreliosis. Survey conducted among 106 general practitioners]', *Presse médicale (Paris, France : 1983).*, 32 1734–6.

McAlister, H., Klementowicz, P., Andrews, C., Fisher, J., Feld, M. and Furman, S. (1989) 'Lyme carditis: An important cause of reversible heart block', *Annals of internal medicine.*, 110 (5): 339–45.

McCord J. M., Fridovich I. (1969) Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). J. Biol. Chem. 244, 6049–6055

Meri, S. (2013) 'Complement activation in diseases presenting with thrombotic microangiopathy', *European journal of internal medicine.*, 24(6), pp. 496–502.

Messenger, A.J. and Barclay, R. (1983) 'Bacteria, iron and pathogenicity', *Biochemical Education*, 11(2), pp. 54–63.

Meyer, J. E. ., Hofnung, M. and Schulz, G. E. (1997) 'Structure of maltoporin from salmonella typhimurium ligated with a nitrophenyl-maltotrioside', *Journal of Molecular Biology*, 266 (4): 761–775.

Midttun, M., Lebech, A., Hansen, K. and Videbaek, J. (1997) 'Lyme carditis: A clinical presentation and long time follow-up', *Scandinavian journal of infectious diseases.*, 29 (2): 153–7.

Miklossy, J. (2008) 'Chronic inflammation and amyloidogenesis in Alzheimer's disease: The emerging role of infection' Journal of Alzheimer's disease., 13(4):381-391.

Miklossy, J., Kasas, S., Zurn, A.D., McCall, S., Yu, S. and McGeer, P.L. (2008) 'Persisting atypical and cystic forms of Borrelia burgdorferi and local inflammation in Lyme neuroborreliosis', *Journal of Neuroinflammation*, 5:40.

Miyamoto, N. and Bakaletz, L. (1996) 'Selective adherence of non-typeable Haemophilus influenzae (NTHi) to mucus or epithelial cells in the chinchilla eustachian tube and middle ear', *Microbial pathogenesis.*, 21 (5): 343–56.

Mizuno, K., Whittaker, M., Bächinger, H. and Whittaker, J. (2004) 'Calorimetric studies on the tight binding metal interactions of Escherichia coli manganese superoxide dismutase', *The Journal of biological chemistry.*, 279 (26): 27339–44.

Moe, G., Tan, S. and Granoff, D. (1999) 'Differences in surface expression of NspA among Neisseria meningitidis group B strains', *Infection and immunity.*, 67 (11): 5664–75.

Moelleken, K. and Hegemann, J.H. (2007) 'The Chlamydia outer membrane protein OmcB is required for adhesion and exhibits biovar-specific differences in glycosaminoglycan binding', *Molecular Microbiology*, 67(2), pp. 403–419.

Montgomery, R., Lusitani, D., Boisfleury, D. and Malawista, S. (2004) 'Tick saliva reduces adherence and area of human neutrophils', *Infection and immunity.*, 72 (5): 2989–94.

Morgan BP. Complement: Clinical Aspects and Relevance to Disease. UK: Academic Press; 1990.

Morona, R., Klose, M. and Henning, U. (1984) 'Escherichia coli K-12 outer membrane protein (OmpA) as a bacteriophage receptor: Analysis of mutant genes expressing altered proteins', *Journal of bacteriology.*, 159 (2): 570–8.

Morrison, T. and Weis, J. (1997) 'Borrelia burgdorferi outer surface protein A (OspA) activates and primes human neutrophils', *Journal of immunology (Baltimore, Md. : 1950).*, 158 (10): 4838–45.

Motaleb, M., Corum, L., Bono, J., Elias, A., Rosa, P., Samuels, D. and Charon, N. (2000) 'Borrelia burgdorferi periplasmic flagella have both skeletal and motility functions', *Proceedings of the National Academy of Sciences of the United States of America.*, 97 (20): 10899–904.

Mulay, V., Caimano, M., Iyer, R., Dunham-Ems, S., Liveris, D., Petzke, M., Schwartz, I. and Radolf, J. (2009) 'Borrelia burgdorferi bba74 is expressed exclusively during tick feeding and is regulated by both arthropod- and mammalian host-specific signals', *Journal of bacteriology.*, 191 (8): 2783–94.

Mulay, V., Caimano, M., Liveris, D., Desrosiers, D., Radolf, J. and Schwartz, I. (2006) 'Borrelia burgdorferi BBA74, a periplasmic protein associated with the outer membrane, lacks porin-like properties', *Journal of bacteriology.*, 189 (5): 2063–8.

Mundodi, V., Kucknoor, A. and Alderete, J. (2007) 'Immunogenic and plasminogen-binding surface-associated alpha-enolase of Trichomonas vaginalis', *Infection and immunity.*, 76(2), pp. 523–31.

Murray, T. and Shapiro, E. (2010) 'Lyme disease', *Clinics in laboratory medicine.*, 30 (1): 311–28.

Mursic, P., Marget, W., Busch, U., Rigler, P. and Hagl, S. (1996) 'Kill kinetics of Borrelia burgdorferi and bacterial findings in relation to the treatment of Lyme borreliosis', *Infection.*, 24 (1): 9–16.

Mygland, A., Ljøstad, U., Fingerle, V., Rupprecht, T., Schmutzhard, E., Steiner, I. and Federation, E. (2009) 'EFNS guidelines on the diagnosis and management of European Lyme neuroborreliosis', *European journal of neurology.*, 17 (1): 8–16.

Nadelman, R., Nowakowski, J., Fish, D., Falco, R., Freeman, K., McKenna, D., Welch, P., Marcus, R., Agüero-Rosenfeld, M., Dennis, D., Wormser, G. and Bite, T. (2001) 'Prophylaxis with single-dose doxycycline for the prevention of Lyme disease after an Ixodes scapularis tick bite', *The New England journal of medicine.*, 345 (2): 79–84.

Nahimana, I., Gern, L., Blanc, D., Praz, G., Francioli, P. and Péter, O. (2004) 'Risk of Borrelia burgdorferi infection in western Switzerland following a tick bite', *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology.*, 23 (8): 603–8.

Naleway, A., Belongia, E., Kazmierczak, J., Greenlee, R. and Davis, J. (2002) 'Lyme disease incidence in Wisconsin: A comparison of state-reported rates and rates from a population-based cohort', *American journal of epidemiology.*, 155(12), pp. 1120–7.

Nandi, B., Nandy, R., Sarkar, A. and Ghose, A. (2005) 'Structural features, properties and regulation of the outer-membrane protein W (OmpW) of Vibrio cholerae', *Microbiology (Reading, England).*, 151 2975–86.

Nigrovic, L. and Thompson, K. (2006) 'The Lyme vaccine: A cautionary tale', *Epidemiology and infection.*, 135 (1): 1–8.

Nishio, M., Okada, N., Miki, T., Haneda, T. and Danbara, H. (2005) 'Identification of the outer-membrane protein PagC required for the serum resistance phenotype in salmonella enterica serovar Choleraesuis', *Microbiology (Reading, England).*, 151, pp. 863–73.

Nocton, J., Bloom, B., Rutledge, B., Persing, D., Logigian, E., Schmid, C. and Steere, A. (1996) 'Detection of Borrelia burgdorferi DNA by polymerase chain reaction in cerebrospinal fluid in Lyme neuroborreliosis', *The Journal of infectious diseases.*, 174 (3): 623–7.

Nocton, J. J., Dressler, F., Rutledge, B. J., Rys, P. N., Persing, D. H. and Steere, A. C. (1994) 'Detection of Borrelia burgdorferi DNA by polymerase chain reaction in Synovial fluid from patients with Lyme arthritis', *New England Journal of Medicine*, 330 (4): 229–234.

Nollert, P., Qiu, H., Caffrey, M., Rosenbusch, J.P., Landau, E.M. (2001) 'Molecular mechanism for the crystallisation of bacteriorhodopsin in lipidic cubic phases', FEBS Lett, 504: 179-186.

Noppa, L., Ostberg, Y., Lavrinovicha, M. and Bergström, S. (2001) 'P13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains', *Infection and immunity.*, 69(5), pp. 3323–34.

Norman, U. M., Moriarty, T. J., Dresser, A. R., Millen, B., Kubes, P. and Chaconas, G. (2008) 'Molecular mechanisms involved in vascular interactions of the Lyme disease Pathogen in a living host', *PLOS Pathog*, 4 (10): 1000169.

Norris, S., Howell, J., Odeh, E., Lin, T., Gao, L. and Edmondson, D. (2010) 'High-throughput plasmid content analysis of Borrelia burgdorferi B31 by using Luminex multiplex technology', *Applied and environmental microbiology.*, 77 (4): 1483–92.

Nowakowski, J., Nadelman, R., Sell, R., McKenna, D., Cavaliere, L., Holmgren, D., Gaidici, A. and Wormser, G. (2003) 'Long-term follow-up of patients with culture-confirmed Lyme disease', *The American journal of medicine.*, 115 (2): 91–6.

Ogden, N., Artsob, H., Lindsay, L. and Sockett, P. (2008) 'Lyme disease: A zoonotic disease of increasing importance to Canadians', *Canadian family physician Médecin de famille canadien.*, 54 (10): 1381–4.

Ojha, S., Sirois, M. and MacInnes, J.I. (2005) 'Identification of Actinobacillus suis genes essential for the colonization of the upper respiratory tract of swine', *Infection and Immunity*, 73(10): 7032–7039.

Okamura, M., Ueda, M., Noda, Y.,. Kuno, Y., Kashimoto, T., Takehara, K., and Nakamura, M. (2012) 'Immunization with outer membrane protein A from Salmonella enterica serovar Enteritidis induces humoral immune response but no protection against homologous challenge in chickens.' Poultry Science. 91 (10): 2444-2449

Önder, Ö., Humphrey, P. T., McOmber, B., Korobova, F., Francella, N., Greenbaum, D. C., Brisson, D. and ozlem@ (2012) 'OspC is potent Plasminogen receptor on surface of Borrelia burgdorferi', *Journal of Biological Chemistry*, 287 (20): 16860–16868.

Oomen, C., Ulsen, van, Gelder, van, Feijen, M., Tommassen, J. and Gros, P. (2004) 'Structure of the translocator domain of a bacterial autotransporter', *The EMBO journal.*, 23 (6): 1257–66.

O'Rourke, M., Traweger, A., Lusa, L., Stupica, D., Maraspin, V., Barrett, P., Strle, F. and Livey, I. (2013) 'Quantitative detection of Borrelia burgdorferi sensu lato in erythema migrans skin lesions using internally controlled duplex real time PCR', *PloS one.*, 8(5), p. e63968.

Oschmann, P., Dorndorf, W., Hornig, C., Schäfer, C., Wellensiek, H. and Pflughaupt, K. (1998) 'Stages and syndromes of neuroborreliosis',*Journal of neurology.*, 245(5): 262–72.

Paetzel M., Dalbey R.E., Strynadka N.C. (2000) 'The structure and mechanism of bacterial type I signal peptidases. A novel antibiotic target', *Pharmacol. Ther*, 87:27–49.

Pal, U., Li, X., Wang, T., Montgomery, R., Ramamoorthi, N., Desilva, A., Bao, F., Yang, X., Pypaert, M., Pradhan, D., Kantor, F., Telford, S., Anderson, J. and Fikrig, E. (2004) 'TROSPA, an Ixodes scapularis receptor for Borrelia burgdorferi', *Cell.*, 119(4), pp. 457–68.

Pal, U., de Silva, A. M., Montgomery, R. R., Fish, D., Anguita, J., Anderson, J. F., Lobet, Y. and Fikrig, E. (2000) 'Attachment of Borrelia burgdorferi within Ixodes scapularis mediated by outer surface protein A', *Journal of Clinical Investigation*, 106 (4): 561–569.

Pal, U., Yang, X., Chen, M., Bockenstedt, L., Anderson, J., Flavell, R., Norgard, M. and Fikrig, E. (2004) 'OspC facilitates Borrelia burgdorferi invasion of Ixodes scapularis salivary glands', *The Journal of clinical investigation.*, 113 (2): 220–30.

Pancholi, V. (2001) 'Multifunctional alpha-enolase: Its role in diseases', *Cellular and molecular life sciences : CMLS.*, 58(7), pp. 902–20.

Pancholi, V. and Fischetti, V. (1998) 'Alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic streptococci', *The Journal of biological chemistry.*, 273(23), pp. 14503–15.

Parveen, N., Cornell, K., Bono, J., Chamberland, C., Rosa, P. and Leong, J. (2006) 'Bgp, a secreted glycosaminoglycan-binding protein of Borrelia burgdorferi strain N40, displays nucleosidase activity and is not essential for infection of immunodeficient mice', *Infection and immunity.*, 74 (5): 3016–20.

Paster, B., Dewhirst, F., Weisburg, W., Tordoff, L., Fraser, G., Hespell, R., Stanton, T., Zablen, L., Mandelco, L. and Woese, C. (1991) 'Phylogenetic analysis of the spirochetes', *Journal of bacteriology.*, 173 (19): 6101–9.

Pausa, M., Pellis, V., Cinco, M., Giulianini, P., Presani, G., Perticarari, S., Murgia, R. and Tedesco, F. (2003) 'Serum-resistant strains of Borrelia burgdorferi evade complement-mediated killing by expressing a CD59-like complement inhibitory molecule', *Journal of immunology (Baltimore, Md. : 1950).*, 170 (6): 3214–22.

Pautsch, A. and Schulz, G. (1998) 'Structure of the outer membrane protein A transmembrane domain', *Nature structural biology.*, 5 (11): 1013–7.

Pautsch, A. and Schulz, G. (2000) 'High-resolution structure of the OmpA membrane domain', *Journal of molecular biology.*, 298 (2): 273–82.

Perego, M., Higgins, C., Pearce, S., Gallagher, M. and Hoch, J. (1991) 'The oligopeptide transport system of Bacillus subtilis plays a role in the initiation of sporulation', *Molecular microbiology.*, 5 (1): 173–85.

Petersen, T.N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) 'SignalP 4.0: discriminating signal peptides from transmembrane regions', *Nature Methods*, 8:785-786.

Petoukhov, M.V. and Svergun, D.I. (2007) 'Analysis of x-ray and neutron scattering from biomacromolecular solutions', *Current Opinion in Structural Biology*, 17(5), pp. 562–571.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, T. E. (2004) 'UCSF Chimera – A visualisation system for exploratory research and analysis', *J Comput Chem*, 13: 1605-1612.

Piesman, J. and Gern, L. (2005) 'Lyme borreliosis in Europe and north America', *Parasitology.*, 129 191–220.

Pitarch, A., Abian, J., Carrascal, M., Sánchez, M., Nombela, C. and Gil, C. (2004) 'Proteomics-based identification of novel Candida albicans antigens for diagnosis of systemic candidiasis in patients with underlying hematological malignancies', *Proteomics.*, 4(10), pp. 3084–106.

Pitarch, A., Jiménez, A., Nombela, C. and Gil, C. (2005) 'Decoding serological response to Candida cell wall immunome into novel diagnostic, prognostic, and therapeutic candidates for systemic candidiasis by proteomic and bioinformatic analyses', *Molecular & cellular proteomics : MCP.*, 5(1), pp. 79–96.

Plante, M., Cadieux, N., Rioux, C., Hamel, J., Brodeur, B. and Martin, D. (1999) 'Antigenic and molecular conservation of the gonococcal NspA protein', *Infection and immunity.*, 67 (6): 2855–61.

Pore, D. and Chakrabarti, M. (2013) 'Outer membrane protein A (OmpA) from Shigella flexneri 2a: A promising subunit vaccine candidate',*Vaccine.*, 31 (36): 3644–50.

Posey, J. and Gherardini, F. (2000) 'Lack of a role for iron in the Lyme disease pathogen', *Science (New York, N.Y.).*, 288 (5471): 1651–3.

Prasadarao, N., Wass, C., Weiser, J., Stins, M., Huang, S. and Kim, K. (1996) 'Outer membrane protein A of Escherichia coli contributes to invasion of brain microvascular endothelial cells', *Infection and immunity.*, 64 (1): 146–53.

Privé, G. (2007) 'Detergents for the stabilization and crystallization of membrane proteins', *Methods*, 41(4), pp. 388–97.

Prince, S., Achtman, M. and Derrick, J. (2002) 'Crystal structure of the OpcA integral membrane adhesin from Neisseria meningitidis', *Proceedings of the National Academy of Sciences of the United States of America.*, 99 (6): 3417–21.

Provencher, S.W. and Glockner, J. (1981) E stimation of globular protein secondary structure from circular dichroism.*Biochemistry* **20**, 33-37.

Pulzova, L., Bhide, M. and Andrej, K. (2009) 'Pathogen translocation across the blood-brain barrier', *FEMS Immunol Med Microbiol.*, 57 (3): 203–213.

Pulzova, L., Kovac, A., Mucha, R., Mlynarcik, P., Bencurova, E., Madar, M., Novak, M. and Bhide, M. (2011) 'OspA-CD40 dyad: Ligand-receptor interaction in the translocation of neuroinvasive Borrelia across the blood-brain barrier', *Scientific Reports*, 1 86.

Pulzova, L. and Bhide, M. (2014) 'Outer surface proteins of Borrelia: Peerless immune evasion tools', *Current protein & peptide science.*, 15 (1): 75–88.

Purser, J., Lawrenz, M., Caimano, M., Howell, J., Radolf, J. and Norris, S. (2003) 'A plasmid-encoded nicotinamidase (PncA) is essential for infectivity of Borrelia burgdorferi in a mammalian host', *Molecular microbiology.*, 48(3), pp. 753–64.

Purser, J. E. and Norris, S. J. (2000) 'Correlation between plasmid content and infectivity in Borrelia burgdorferi', *Proceedings of the National Academy of Sciences*, 97 (25): 13865–13870.

Putnam, C., Hammel, M., Hura, G. and Tainer, J. (2007) 'X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution', *Quarterly reviews of biophysics.*, 40(3), pp. 191–285.

Qiu, H. and Caffrey, M. (2000) 'The phase diagram of the monoolein/water system: metastability and equilibrium aspects', *Biomaterials.*, 21(3): 223-234.

Radolf, J., Bourell, K., Akins, D., Brusca, J. and Norgard, M. (1994) 'Analysis of Borrelia burgdorferi membrane architecture by freeze-fracture electron microscopy', *Journal of bacteriology.*, 176 (1): 21–31.

Radolf, J., Caimano, M., Stevenson, B. and Hu, L. (2012) 'Of ticks, mice and men: Understanding the dual-host lifestyle of Lyme disease spirochaetes', *Nature reviews. Microbiology.*, 10 (2): 87–99.

Ramamoorthi, N., Narasimhan, S., Pal, U., Bao, F., Yang, X., Fish, D., Anguita, J., Norgard, M., Kantor, F., Anderson, J., Koski, R. and Fikrig, E. (2005) 'The Lyme disease agent exploits a tick protein to infect the mammalian host', *Nature.*, 436 (7050): 573–7.

Rambo, R.P. and Tainer, J.A. (2013) 'Accurate assessment of mass, models and resolution by small-angle scattering', Nature, 496(7446), pp. 477–481.

Rao KV, Cullen WP. (1960) Streptonigrin, an antitumor substance. I. Isolation and characterization. Antibiot Annu. 7950-7953.

Rasmussen, S., DeVree, B., Zou, Y., Kruse, A., Chung, K., Kobilka, T., Thian, F., Chae, P., Pardon, E., Calinski, D., Mathiesen, J., Shah, S., Lyons, J., Caffrey, M., Gellman, S., Steyaert, J., Skiniotis, G., Weis, W., Sunahara, R. and Kobilka, B. (2011) 'Crystal structure of the β2 adrenergic receptor-gs protein complex', Nature., 477(7366), pp. 549–55.

Ravdin, L., Hilton, E., Primeau, M., Clements, C. and Barr, W. (1996) 'Memory functioning in Lyme borreliosis', *The Journal of clinical psychiatry.*, 57 (7): 282–6.

Reik, L., Steere, A.C., Bartenhagen, N.H., Shope, R.E., Malawista, S.E. (1979) 'Neurologic abnormalities of Lyme disease', Med Baltim., (58): 281-294.

Revel, A., Blevins, J., Almazán, C., Neil, L., Kocan, K., de, la, Hagman, K. and Norgard, M. (2005) 'BptA (bbe16) is essential for the persistence of the Lyme disease spirochete, Borrelia burgdorferi, in its natural tick vector', *Proceedings of the National Academy of Sciences of the United States of America.*, 102 (19): 6972–7.

Reyes-Lamothe, R., Nicolas, E. and Sherratt, D. (2012) 'Chromosome replication and segregation in bacteria', *Annual review of genetics.*, 46 121–43.

Ribeiro, J. (1987) 'Role of saliva in blood-feeding by arthropods', *Annual review of entomology.*, 32 463–78.

Ribeiro, J., Alarcon-Chaidez, F., Francischetti, I., Mans, B., Mather, T., Valenzuela, J. and Wikel, S. (2006) 'An annotated catalog of salivary gland transcripts from Ixodes scapularis ticks', *Insect biochemistry and molecular biology.*, 36 (2): 111–29.

Ristow, P., Bourhy, P., McBride, F.W. da C., Figueira, C.P., Huerre, M., Ave, P., Girons, I.S., Ko, A.I. and Picardeau, M. (2007) 'The OmpA-Like protein Loa22 is essential for Leptospiral virulence', *PLoS Pathogens*, 3 (7): e97.

Rizzoli, A., Hauffe, H., Carpi, G., Vourc, H., Neteler, M. and Rosa, R. (2011) 'Lyme borreliosis in Europe', *Euro surveillance : bulletin Européen sur les maladies transmissibles = European communicable disease bulletin.*, 16 (27) .

Robert, V., Volokhina, E., Senf, F., Bos, M., Gelder, V. and Tommassen, J. (2006) 'Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific c-terminal motif', *PLoS biology.*, 4(11).

Rosa, P. (2005) 'Lyme disease agent borrows a practical coat', *Nature medicine.*, 11 (8): 831–2.

Rosato, R., Fernandez, R., Paz, L., Singh, C. and Rosato, A. (2014) 'TCA cycle-mediated generation of ROS is a key mediator for HeR-MRSA survival under β-lactam antibiotic exposure', PloS one., 9(6).

Rudenko, N., Golovchenko, M., Grubhoffer, L. and Oliver, J. (2010) 'Borrelia carolinensis sp. Nov., a novel species of the Borrelia burgdorferi sensu lato complex isolated from rodents and a tick from the south-eastern USA', *International journal of systematic and evolutionary microbiology.*, 61 381–3.

Rupprecht, T., Koedel, U., Heimerl, C., Fingerle, V., Paul, R., Wilske, B. and Pfister, H. (2006) ‘Adhesion of Borrelia garinii to neuronal cells is mediated by the interaction of OspA with proteoglycans’, *Journal of neuroimmunology.*, 175, pp. 5–11.

Russell, T. and Johnson, B. (2013) ‘Lyme disease spirochaetes possess an aggrecan-binding protease with aggrecanase activity’, *Molecular microbiology.*, 90(2), pp. 228–40.

Rutten, L., Geurtsen, J., Lambert, W., Smolenaers, J., Bonvin, A., Haan, de, der, van, Egmond, Gros, P. and Tommassen, J. (2006) ‘Crystal structure and catalytic mechanism of the LPS 3-O-deacylase PagL from Pseudomonas aeruginosa’, *Proceedings of the National Academy of Sciences of the United States of America.*, 103(18), pp. 7071–6.

Sarkar, P.K. and Doty, P. (1966) ‘The optical rotatory properties of the beta-configuration in polypeptides and proteins’, *PNAS*, 55(4), pp. 981–989.

Salazar, J., Gerber, M. and Goff, C. (1993) 'Long-term outcome of Lyme disease in children given early treatment', *The Journal of pediatrics.*, 122 (4): 591–3.

Salaün, C., James, D. and Chamberlain, L. (2004) 'Lipid rafts and the regulation of exocytosis', *Traffic (Copenhagen, Denmark).*, 5 (4): 255–64.

Sali, A. & Blundell, T.L. (1993) 'Comparative protein modelling by satisfaction of spatial restraints'. J. Mol. Biol. 234, 779-815.

Sanchez, J. (2015) ‘Clinical manifestations and treatment of Lyme disease’, *Clinics in Laboratory Medicine*, 35 (4): 765-778.

Sarkar, P.K. and Doty, P. (1966) ‘The optical rotatory properties of the beta-configuration in polypeptides and proteins’, *PNAS*, 55(4), pp. 981–989.

Sarma, J. and Ward, P. (2010) 'The complement system', *Cell and tissue research.*, 343 (1): 227–35.

Sato, S. and Nakazawa, K. (1978) ‘Purification and properties of superoxide dismutase from Thermus thermophilus HB8’, *Journal of biochemistry.*, 83(4), pp. 1165–71.

Schifferli JA, Ng YC, Peters DK. (1986) The role of complement and its receptor in the elimination of immune complexes. *N Engl J Med*. 21;315(8):488–495.

Schirmer, T., Keller, T., Wang, Y. and Rosenbusch, J. (1995) 'Structural basis for sugar translocation through maltoporin channels at 3.1 A resolution', *Science (New York, N.Y.).*, 267 (5197): 512–4.

Schmidt CQ, Herbert AP, Hocking HG, Uhrin D, Barlow PN. Translational Mini-Review Series on Complement Factor H: Structural and functional correlations for factor H. Clinical & Experimental Immunology. 2008;151:14–24.

Schneidman-Duhovny, D., Hammel, M. and Sali, A. (2010) ‘FoXS: A web server for rapid computation and fitting of SAXS profiles’, Nucleic Acids Research, 38(Web Server), pp. 540–544.

Schroder, N. W. J., Schombel, U., Heine, H., Gobel, U. B., Zahringer, U. and Schumann, R. R. (2003) 'Acylated Cholesteryl Galactoside as a novel Immunogenic motif in Borrelia burgdorferi Sensu Stricto', *Journal of Biological Chemistry*, 278 (36): 33645–33653.

Schröder, N., Eckert, J., Stübs, G. and Schumann, R. (2008) 'Immune responses induced by spirochetal outer membrane lipoproteins and glycolipids', *Immunobiology.*, 213 329–40.

Schuijt, T. J., Coumou, J., Narasimhan, S., Dai, J., DePonte, K., Wouters, D., Brouwer, M., Oei, A., Roelofs, J. J. T. H., van Dam, A. P., van der Poll, T., van't Veer, C., Hovius, J. W. and Fikrig, E. (2011) 'A tick Mannose-Binding Lectin inhibitor interferes with the vertebrate complement cascade to enhance transmission of the Lyme disease agent', *Cell Host & Microbe*, 10 (2): 136–146.

Schulz, G. (2000) 'Beta-barrel membrane proteins', *Current opinion in structural biology.*, 10 (4): 443–7.

Schutzer, S., Coyle, P., Krupp, L., Deng, Z., Belman, A., Dattwyler, R. and Luft, B. (1997) 'Simultaneous expression of Borrelia OspA and OspC and IgM response in cerebrospinal fluid in early neurologic Lyme disease', *The Journal of clinical investigation.*, 100 (4): 763–7.

Schutzer, S., Fraser-Liggett, C., Casjens, S., Qiu, W., Dunn, J., Mongodin, E. and Luft, B. (2010) 'Whole-genome sequences of thirteen isolates of Borrelia burgdorferi', *Journal of bacteriology.*, 193 (4): 1018–20.

Schwab, J., Hammerschmidt, C., Richter, D., Skerka, C., Matuschka, F.-R., Wallich, R., Zipfel, P. F. and Kraiczy, P. (2013) 'Borrelia valaisiana resist complement-mediated killing independently of the recruitment of immune regulators and inactivation of complement components', *PLoS One*, 8 (1): e53659.

Schwan, T., Policastro, P., Miller, Z., Thompson, R., Damrow, T. and Keirans, J. (2003) 'Tick-borne relapsing fever caused by Borrelia hermsii, Montana', *Emerging infectious diseases.*, 9 (9): 1151–4.

Schwan, T. G. and Piesman, J. (2000) 'Temporal changes in outer surface proteins A and C of the Lyme disease-associated Spirochete, Borrelia burgdorferi, during the chain of infection in ticks and mice', *Journal of clinical microbiology*, 38 (1): 382–388.

Schwan, T. and Hinnebusch, B. (1998) 'Bloodstream- versus tick-associated variants of a relapsing fever bacterium', *Science (New York, N.Y.).*, 280 (5371): 1938–40.

Serino, L., Nesta, B., Leuzzi, R., Fontana, M., Monaci, E., Mocca, B., Cartocci, E., Masignani, V., Jerse, A., Rappuoli, R. and Pizza, M. (2007) 'Identification of a new OmpA-like protein in Neisseria gonorrhoeae involved in the binding to human epithelial cells and in vivo colonization', *Molecular microbiology.*, 64 (5): 1391–403.

Sethi, N., Sondey, M., Bai, Y., Kim, K. and Cadavid, D. (2006) 'Interaction of a neurotropic strain of Borrelia turicatae with the cerebral microcirculation system', *Infection and immunity.*, 74 (11): 6408–18.

Setubal, J. C., Reis, M., Matsunaga, J. and Haake, D. A. (2006) 'Lipoprotein computational prediction in spirochaetal genomes', Microbiology, 152: 113-121.

Shadick, N., Phillips, C., Logigian, E., Steere, A., Kaplan, R., Berardi, V., Duray, P., Larson, M., Wright, E., Ginsburg, K., Katz, J. and Liang, M. (1994) 'The long-term clinical outcomes of Lyme disease. A population-based retrospective cohort study', *Annals of internal medicine.*, 121 (8): 560–7.

Shapiro, E., Gerber, M., Holabird, N., Berg, A., Feder, H., Bell, G., Rys, P. and Persing, D. (1992) 'A controlled trial of antimicrobial prophylaxis for Lyme disease after deer-tick bites', *The New England journal of medicine.*, 327 (25): 1769–73.

Shimizu, K. (2013) 'Regulation systems of bacteria such as Escherichia coli in response to nutrient limitation and environmental stresses', Metabolites, 4(1), pp. 1–35.

Shin, S., Lu, G., Cai, M. and Kim, K. (2005) 'Escherichia coli outer membrane protein A adheres to human brain microvascular endothelial cells', *Biochemical and biophysical research communications.*, 330 (4): 1199–204.

Shoberg, R. and Thomas, D. (1993) 'Specific adherence of Borrelia burgdorferi extracellular vesicles to human endothelial cells in culture', *Infection and immunity.*, 61(9), pp. 3892–900.

Siegel, C., Hallström, T., Skerka, C., Eberhardt, H., Uzonyi, B., Beckhaus, T., Karas, M., Wallich, R., Stevenson, B., Zipfel, P. and Kraiczy, P. (2010) 'Complement factor h-related proteins CFHR2 and CFHR5 represent novel ligands for the infection-associated CRASP proteins of Borrelia burgdorferi', *PloS one.*, 5 (10): e13519.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G. and California (2011) 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', EMBO Report, 7(1), p. 539

Singh, R., Shasany, A., Aggarwal, A., Sinha, S., Sisodia, B., Khanuja, S. and Misra, R. (2007) 'Low molecular weight proteins of outer membrane of salmonella typhimurium are immunogenic in salmonella induced reactive arthritis revealed by proteomics', *Clinical and experimental immunology.*, 148 (3): 486–93.

Skare, J. T., Mirzabekov, T. A., Shang, E. S., Blanco, D. R., Erdjument-Bromage, H., Bunikis, J., Bergström, S., Tempst, P., Kagan, B. L., Miller, J. N. and Lovett, M. A. (1997) 'The Oms66 (p66) protein is a Borrelia burgdorferi porin', *Infection and immunity*, 65 (9): 3654–3661.

Smialowski, P., Schmidt, T., Cox, J., Kirschner, A. and Frishman, D. (2005) 'Will my protein crystallize? A sequence-based predictor', *Proteins.*, 62(2), pp. 343–55.

Smith, R., O'Connell, S., Palmer, S. (2000) 'Lyme disease surveillance in England and Wales 1986-1998', *Emerging Infectious Dieases*, 6 (4):401-407

Smith, R., Schoen, R., Rahn, D., Sikand, V., Nowakowski, J., Parenti, D., Holman, M., Persing, D. and Steere, A. (2002) 'Clinical characteristics and treatment outcome of early Lyme disease in patients with microbiologically confirmed erythema migrans', *Annals of internal medicine.*, 136 (6): 421–8.

Smith, S., Mahon, V., Lambert, M. and Fagan, R. (2007) 'A molecular Swiss army knife: OmpA structure, function and expression', *FEMS microbiology letters.*, 273 (1): 1–11.

Snäll, J., Linnér, A., Uhlmann, J., Siemens, N., Ibold, H., Janos, M., Linder, A., Kreikemeyer, B., Herwald, H., Johansson, L. and Norrby-Teglund, A. (2016) 'Differential neutrophil responses to bacterial stimuli: Streptococcal strains are potent inducers of heparin-binding protein and resistin-release', Scientific reports., 6, p. 21288.

Snijder, H., Ubarretxena-Belandia, I., Blaauw, M., Kalk, K., Verheij, H., Egmond, M., Dekker, N. and Dijkstra, B. (1999) 'Structural evidence for dimerization-regulated activation of an integral membrane phospholipase', *Nature.*, 401 (6754): 717–21.

Sood, S., Salzman, M., Johnson, B., Happ, C., Feig, K., Carmody, L., Rubin, L., Hilton, E. and Piesman, J. (1997) 'Duration of tick attachment as a predictor of the risk of Lyme disease in an area in which Lyme disease is endemic', *The Journal of infectious diseases.*, 175 (4): 996–9.

Sreerema, N. and Woody, R.W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal. Biochem.* **209**, 32-44.

Sreerama, N. and Woody, R. (2000) 'Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set',*Analytical biochemistry.*, 287(2), pp. 252–60.

Srivastava, A., Yano, J., Hirozane, Y., Kefala, G., Gruswitz, F., Snell, G., Lane, W., Ivetac, A., Aertgeerts, K., Nguyen, J., Jennings, A. and Okada, K. (2014) 'High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875', Nature., 513(7516), pp. 124–7.

Stafford, K., Denicola, A., Pound, J., Miller, J. and George, J. (2009) 'Topical treatment of white-tailed deer with an acaricide for the control of Ixodes scapularis (Acari: Ixodidae) in a Connecticut Lyme borreliosis hyperendemic community', *Vector borne and zoonotic diseases (Larchmont, N.Y.).*, 9 (4): 371–9.

Stanek, G. and Kahl, O. (2000) 'Chemoprophylaxis for Lyme borreliosis?', *Zentralblatt für Bakteriologie : international journal of medical microbiology.*, 289 655–65.

Stanek, G., Wormser, G., Gray, J. and Strle, F. (2011) 'Lyme borreliosis', *Lancet (London, England).*, 379 (9814): 461–73.

Steere, A., Batsford, W., Weinberg, M., Alexander, J., Berger, H., Wolfson, S. and Malawista, S. (1980) 'Lyme carditis: Cardiac abnormalities of Lyme disease', *Annals of internal medicine.*, 93 (1): 8–16.

Steere, A., Coburn, J. and Glickstein, L. (2004) 'The emergence of Lyme disease', *The Journal of clinical investigation.*, 113 (8): 1093–101.

Steere, A., Hardin, J. and Malawista, S. (1978) 'Lyme arthritis: A new clinical entity', *Hospital practice.*, 13 (4): 143–58.

Steere, A., Malawista, S., Snydman, D., Shope, R., Andiman, W. and Steele, F. (1977) 'Lyme arthritis: An epidemic of oligoarticular arthritis in children and adults in three connecticut communities', *Arthritis and rheumatism.*, 20 (1): 7–17.

Steere, A., McHugh, G., Damle, N. and Sikand, V. (2008) 'Prospective study of serologic tests for lyme disease', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.*, 47 (2): 188–95.

Steere, A., Schoen, R. and Taylor, E. (1987) 'The clinical evolution of Lyme arthritis', *Annals of internal medicine.*, 107 (5): 725–31.

Steere, A., Sikand, V., Meurice, F., Parenti, D., Fikrig, E., Schoen, R., Nowakowski, J., Schmid, C., Laukamp, S., Buscarino, C. and Krause, D. (1998) 'Vaccination against Lyme disease with recombinant Borrelia burgdorferi outer-surface lipoprotein A with adjuvant. Lyme disease vaccine study group', *The New England journal of medicine.*, 339 (4): 209–15.

Steere, A. and Angelis, S. (2006) 'Therapy for Lyme arthritis: Strategies for the treatment of antibiotic-refractory arthritis', *Arthritis and rheumatism.*, 54 (10): 3079–86.

Stewart, P., Byram, R., Grimm, D., Tilly, K. and Rosa, P. (2005) 'The plasmids of Borrelia burgdorferi: Essential genetic elements of a pathogen', *Plasmid.*, 53 (1): 1–13.

Stewart, P., Wang, X., Bueschel, D., Clifton, D., Grimm, D., Tilly, K., Carroll, J., Weis, J. and Rosa, P. (2006) 'Delineating the requirement for the Borrelia burgdorferi virulence factor OspC in the mammalian host', *Infection and immunity.*, 74 (6): 3547–53.

Stewart, P., Wang, X., Bueschel, D., Clifton, D., Grimm, D., Tilly, K., Carroll, J., Weis, J. and Rosa, P. (2006) 'Delineating the requirement for the Borrelia burgdorferi virulence factor OspC in the mammalian host', *Infection and immunity.*, 74 (6): 3547–53.

Stocker, B., Nurminen, M. and Mäkelä, P. (1979) 'Mutants defective in the 33K outer membrane protein of salmonella typhimurium', *Journal of bacteriology.*, 139 (2): 376–83.

Sujiwattanarat, P., Pongsanarakul, P., Temsiripong, Y., Thawornkuno, C., Uno, Y., Unajak, S., Matsuda, Y., Choowongkomon, K. and Srikulnath, K. (2015) 'Molecular cloning and characterization of Siamese crocodile (Crocodylus siamensis) copper, zinc superoxide dismutase (CSI-Cu, Zn-SOD) gene',*Comparative biochemistry and physiology. Part A, Molecular & integrative physiology.*, 191, pp. 187–95.

Strle, F., Nelson, J., Ruzic-Sabljic, E., Cimperman, J., Maraspin, V., Lotric-Furlan, S., Cheng, Y., Picken, M., Trenholme, G. and Picken, R. (1996) 'European Lyme borreliosis: 231 culture-confirmed cases involving patients with erythema migrans', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.*, 23 (1): 61–5.

Strle, F., Pleterski-Rigler, D., Stanek, G., Pejovnik-Pustinek, A., Ruzic, E. and Cimperman, J. (1992) 'Solitary borrelial lymphocytoma: Report of 36 cases', *Infection.*, 20 (4): 201–6.

Strle, F., Ruzić-Sabljić, E., Cimperman, J., Lotric-Furlan, S. and Maraspin, V. (2006) 'Comparison of findings for patients with Borrelia garinii and Borrelia afzelii isolated from cerebrospinal fluid', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.*, 43 (6): 704–10.

Stupica, D., Lusa, L., Ruzić-Sabljić, E., Cerar, T. and Strle, F. (2012) 'Treatment of erythema migrans with doxycycline for 10 days versus 15 days', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.*, 55 (3): 343–50.

Struyve M., Moons M., Tommassen J. (1991) 'Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein', *J. Mol. Biol,.* 218:141–148.

Stübs, G., Fingerle, V., Wilske, B., Göbel, U., Zähringer, U., Schumann, R. and Schröder, N. (2009) 'Acylated cholesteryl galactosides are specific antigens of borrelia causing lyme disease and frequently induce antibodies in late stages of disease', *The Journal of biological chemistry.*, 284 (20): 13326–34.

Su, H., Raymond, L., Rockey, D.D., Fischer, E., Hackstadt, T. and Caldwell, H.D. (1996) 'A recombinant Chlamydia trachomatis major outer membrane protein binds to heparan sulfate receptors on epithelial cells', *PNAS*, 93 (20): 403–419.

Sullivan, T., LaScolea, L. and Neter, E. (1982) 'Relationship between the magnitude of bacteremia in children and the clinical disease', *Pediatrics.*, 69 (6): 699–702.

Sutcliffe, I. and Russell, R. (1995) 'Lipoproteins of gram-positive bacteria', *Journal of bacteriology.*, 177 (5): 1123–8.

Svergun D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. 25, 495-503.

Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J.* 2879-2886.

Svergun, D.I. (2007) 'Small-angle scattering studies of macromolecular solutions', *Journal of Applied Crystallography*, 40(s1):10-17.

Svergun, D.I., Petoukhov, M.V. and Koch, M.H.J. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.*, **80**, 2946-2953.

Szczepanski, A. and Benach, J.L. (1991) 'Lyme borreliosis: Host responses to Borrelia burgdorferi',*Microbiological Reviews*, 55(1), pp. 21–34.

Szczepanski, A., Furie, M. B., Benach, J. L., Lane, B. P. and Fleit, H. B. (1990) 'Interaction between Borrelia burgdorferi and endothelium in vitro', *The Journal of Clinical Investigation*, 85 (5): 1637–1647.

Taeye, de, Kreuk, L., Dam, van, Hovius, J. and Schuijt, T. (2013) 'Complement evasion by Borrelia burgdorferi: It takes three to tango', *Trends in parasitology.*, 29(3), pp. 119–28.

Takayama, K., Rothenberg, R.J. and Barbour, A.G. (1987) 'Absence of lipopolysaccharide in the Lyme disease spirochete, Borrelia burgdorferi', *Infection and Immunity*, 55(9), pp. 2311–2313.

Tamm, L.K., Hong, H. and Liang, B. (2004) 'Folding and assembly of β-barrel membrane proteins', Biochimica et Biophysica Acta (BBA) - Biomembranes, 1666(s 1–2), pp. 250–263.

Tegla, C. A., Cudrici, C., Patel, S., Trippe, R., Rus, V., Niculescu, F., & Rus, H. (2011). Membrane attack by complement: the assembly and biology of terminal complement complexes. *Immunologic Research*, *51*(1), 45–60.

Tegoni, M., Ramoni, R., Bignetti, E., Spinelli, S. and Cambillau, C. (1996) 'Domain swapping creates a third putative combining site in bovine odorant binding protein dimer', *Nature structural biology.*, 3(10), pp. 863–7.

The New Yorker. 2012. Mitt Romney versus Lyme disease and science. http://www.newyorker.com/news/news-desk/mitt-romney-versus-lyme-disease-and-science.

The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC

Thein, M., Bonde, M., Bunikis, I., Denker, K., Sickmann, A., Bergström, S. and Benz, R. (2012) 'DipA, a pore-forming protein in the outer membrane of Lyme disease spirochetes exhibits specificity for the permeation of dicarboxylates', *PloS one.*, 7(5).

Thomas, D. D. and Comstock, L. E. (1989) 'Interaction of Lyme disease spirochetes with cultured eucaryotic cells', *Infection and Immunity*, 57 (4): 1324–1326.

Thomas, V., Kinkead, L., Janssen, A., Schaeffer, C., Woods, K., Lindgren, J., Peaster, J., Chaudhari, S., Sadykov, M., Jones, J., AbdelGhani, S., Zimmerman, M., Bayles, K., Somerville, G. and Fey, P. (2013) 'A dysfunctional tricarboxylic acid cycle enhances fitness of staphylococcus epidermidis during β-lactam stress', mBio., 4(4).

Tibbles, C. and Edlow, J. (2007) 'Does this patient have erythema migrans?', *JAMA.*, 297 (23): 2617–27.

Tiefenbrunn, T., Liu, W., Chen, Y., Katritch, V., Stout, C., Fee, J. and Cherezov, V. (2011) 'High resolution structure of the ba3 cytochrome c oxidase from Thermus thermophilus in a lipidic environment', PloS one., 6(7), pp. 686–690.

Tilly, K., Casjens, S., Stevenson, B., Bono, J., Samuels, D., Hogan, D. and Rosa, P. (1997) 'The Borrelia burgdorferi circular plasmid cp26: Conservation of plasmid structure and targeted inactivation of the ospC gene', *Molecular microbiology.*, 25 (2): 361–73.

Tokunaga, M. and Wu, H. (1982) 'Post-translational modification and processing of Escherichia coli prolipoprotein in vitro', *Proceedings of the National Academy of Sciences of the United States of America.*, 79 (7): 2255–9.

Toledo, A., Coleman, J., Kuhlow, C., Crowley, J. and Benach, J. (2011) 'The enolase of Borrelia burgdorferi is a plasminogen receptor released in outer membrane vesicles', *Infection and immunity.*, 80 (1): 359–68.

Tommassen, J. (2010) 'Assembly of outer-membrane proteins in bacteria and mitochondria', *Microbiology (Reading, England).*, 156, pp. 2587–96.

Touw, D.S., Patel, D.R. and van den Berg, B. (2010) 'The crystal structure of OprG from Pseudomonas aeruginosa, a potential channel for transport of hydrophobic molecules across the outer membrane', *PLoS ONE*, 5(11), p. e15016.

Troxell, B., Xu, H. and Yang, X. (2012) 'Borrelia burgdorferi, a pathogen that lacks iron, encodes manganese-dependent superoxide dismutase essential for resistance to streptonigrin', *The Journal of biological chemistry.*, 287 (23): 19284–93.

Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R. and Yoshikawa, S. (1996) 'The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8A', Science., 272 (5265): 1136-1144.

Tveitnes, D., Øymar, K. and Natås, O. (2009) 'Laboratory data in children with Lyme neuroborreliosis, relation to clinical presentation and duration of symptoms', *Scandinavian journal of infectious diseases.*, 41 (5): 355–62.

Tyson, K., Elkins, C., Patterson, H., Fikrig, E. and Silva, de (2007) 'Biochemical and functional characterization of Salp20, an Ixodes scapularis tick salivary protein that inhibits the complement pathway', *Insect molecular biology.*, 16 (4): 469–79.

Ulmschneider, M. and Sansom, M. (2001) 'Amino acid distributions in integral membrane protein structures', Biochimica et biophysica acta., 1512(1), pp. 1–14.

Van Stokkum, I.H.M., Spoelder, H.J.W., Bloemendal, M., Van Grondelle, R., and Groen, F.C.A. (1990) Estimation of protein secondary structure and error analysis from CD spectra. *Anal. Biochem.* **191**, 110-118.

Vance, C. and Miller, A. (1998) 'Spectroscopic comparisons of the pH dependencies of Fe-substituted (Mn)superoxide dismutase and Fe-superoxide dismutase', *Biochemistry.*, 37 (16): 5518–27.

Vandeputte-Rutten, L., Bos, M., Tommassen, J. and Gros, P. (2003) 'Crystal structure of Neisserial surface protein A (NspA), a conserved outer membrane protein with vaccine potential', *The Journal of biological chemistry.*, 278 (27): 24825–30.

Vandeputte-Rutten, L., Kramer, R., Kroon, J., Dekker, N., Egmond, M. and Gros, P. (2001) 'Crystal structure of the outer membrane protease OmpT from Escherichia coli suggests a novel catalytic site', *The EMBO journal.*, 20 (18): 5033–9.

Verhoeven, G.S., Dogterom, M. and Blaauwen, T. den (2013) 'Absence of long-range diffusion of OmpA in E. Coli is not caused by its peptidoglycan binding domain', BMC Microbiology, 13(66).

Verma, A., Brissette, C., Bowman, A. and Stevenson, B. (2009) 'Borrelia burgdorferi BmpA is a laminin-binding protein', *Infection and immunity.*, 77 (11): 4940–6.

Vidal-Ingigliardi, D., Lewenza, S. and Buddelmeijer, N. (2007) 'Identification of essential residues in apolipoprotein N-acyl transferase, a member of the CN hydrolase family', *Journal of bacteriology.*, 189 (12): 4456–64.

Vincentelli, R., Canaan, S., Campanacci, V., Valencia, C., Maurin, D., Frassinetti, F., Scappucini-Calvo, L., Bourne, Y., Cambillau, C. and Bignon, C. (2004) 'High-throughput automated refolding screening of inclusion bodies', 13(10), pp. 2782–2792.

Vogel, H. and Jähnig, F. (1986) 'Models for the structure of outer-membrane proteins of Escherichia coli derived from raman spectroscopy and prediction methods', *Journal of molecular biology.*, 190 (2): 191–9.

Vogt, J. and Schulz, G. (1999) 'The structure of the outer membrane protein OmpX from Escherichia coli reveals possible mechanisms of virulence', *Structure (London, England : 1993).*, 7 (10): 1301–9.

Volkov, V. & Svergun, D.I. (2003). Uniqueness of ab-initio shape determination in small-angle scattering. J. Appl. Cryst. 36, 860-864.

Vrethem, M., Hellblom, L., Widlund, M., Ahl, M., Danielsson, O., Ernerudh, J. and Forsberg, P. (2002) 'Chronic symptoms are common in patients with neuroborreliosis -- a questionnaire follow-up study', *Acta neurologica Scandinavica.*, 106(4), pp. 205–8.

Wallis R. (2007) 'Interactions between mannose-binding lectin and MASPs during complement activation by the lectin pathway.' *Immunobiology.* 212(4–5):289–99.

Walport, M. (2001) 'Complement. First of two parts', *The New England journal of medicine.*, 344(14), pp. 1058–66.

Wang, P., Dadhwal, P., Cheng, Z., Zianni, M. R., Rikihisa, Y., Liang, F. T., & Li, X. (2013). *Borrelia burgdorferi* Oxidative Stress Regulator BosR Directly Represses Lipoproteins Primarily Expressed in the Tick during Mammalian Infection. *Molecular Microbiology*, *89*(6), 1140–1153.

Wang, Q., Kaan, H., Hooda, R., Goh, S. and Sondermann, H. (2008) 'Structure and plasticity of Endophilin and sorting Nexin 9', *Structure (London, England : 1993).*, 16(10), pp. 1574–87.

Wang, Y. and Kim, K. (2002) 'Role of OmpA and IbeB in Escherichia coli K1 invasion of brain microvascular endothelial cells in vitro and in vivo', *Pediatric research.*, 51(5), pp. 559–63.

Wang, Y., Xu, W. and Chitnis, P.R. (2009) 'Identification and bioinformatic analysis of the membrane proteins of synechocystis sp. PCC 6803', Proteome Science, 7(1), p. 11.

Wass MN, Kelley LA & Sternberg MJ (2010) 3DLigandSite: predicting ligand-binding sites using similar structures. *NAR* 38 Suppl:W469-73

Webb, B., Sali, A. (2014) 'Comparative Protein Structure Modeling Using Modeller'. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., 5.6.1-5.6.32.

Weber, G. (1983) 'Stability of oligomeric proteins and its bearing on their association equilibria (a reply)', PNAS, 80(17), pp. 5303–5304.

Weber, K., Preac-Mursic, V., Wilske, B., Thurmayr, R., Neubert, U. and Scherwitz, C. (1990) 'A randomized trial of ceftriaxone versus oral penicillin for the treatment of early European Lyme borreliosis', *Infection.*, 18 (2): 91–6.

Weiser, J. and Gotschlich, E. (1991) 'Outer membrane protein A (OmpA) contributes to serum resistance and pathogenicity of Escherichia coli K-1', *Infection and immunity.*, 59 (7): 2252–8.

Weiss, M. and Schulz, G. (1992) 'Structure of porin refined at 1.8 A resolution', *Journal of molecular biology.*, 227 (2): 493–509.

Whittaker, M. and Whittaker, J. (2000) 'Recombinant superoxide dismutase from a hyperthermophilic archaeon, Pyrobaculum aerophilium', *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry.*, 5(3), pp. 402–8.

White H. L., White J. R. (1968) Lethal action and metabolic effects of streptonigrin on Escherichia coli. Mol. Pharmacol. 4, 549–565

Whitehouse, C., Williams, L. and Austin, F. (1997) 'Identification of superoxide dismutase activity in Borrelia burgdorferi', *Infection and immunity.*, 65 (11): 4865–8.

Whitmore, L. and Wallace, B. (2004) 'DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data', *Nucleic acids research.*, 32, pp. 668–673.

Whitmore, L. and Wallace, B. (2007) 'Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases', *Biopolymers.*, 89(5), pp. 392–400.

WHO (2013) 'The world health report 2004 - changing history', available at http://www.who.int/whr/2004/en.

Wood, E., Tamborero, S., Mingarro, I. and Esteve-Gassent, M. (2013) 'BB0172, a Borrelia burgdorferi outer membrane protein that binds integrin α3β1', *Journal of bacteriology.*, 195(15), pp. 3320–30.

Wormser, G., Dattwyler, R., Shapiro, E., Halperin, J., Steere, A., Klempner, M., Krause, P., Bakken, J., Strle, F., Stanek, G., Bockenstedt, L., Fish, D., Dumler, J. and Nadelman, R. (2006) 'The clinical assessment, treatment, and prevention of lyme disease, human granulocytic anaplasmosis, and babesiosis: Clinical practice guidelines by the infectious diseases society of America', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.*, 43 (9): 1089–134.

Wormser, G., Ramanathan, R., Nowakowski, J., McKenna, D., Holmgren, D., Visintainer, P., Dornbush, R., Singh, B. and Nadelman, R. (2003) 'Duration of antibiotic therapy for early Lyme disease. A randomized, double-blind, placebo-controlled trial', *Annals of internal medicine.*, 138 (9): 697–704.

Wright, D. (2009) 'Borrel's accidental legacy', *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases.*, 15 (5): 397–9.

Wu, X.-B., Tian, L.-H., Zou, H.-J., Wang, C.-Y., Yu, Z.-Q., Tang, C.-H., Zhao, F.-K. and Pan, J.-Y. (2013) 'Outer membrane protein OmpW of Escherichia coli is required for resistance to phagocytosis', *Research in Microbiology*, 164 (8): 848–855.

Xie, Y. and Wetlaufer, D.B. (1996) 'Control of aggregation in protein refolding: The temperature-leap tactic', Protein Science, 5(3), pp. 517–523.

Xu, C., Wang, S., Ren, H., Lin, X., Wu, L. and Peng, X. (2005) 'Proteomic analysis on the expression of outer membrane proteins of Vibrio alginolyticus at different sodium concentrations', *Proteomics.*, 5 (12): 3142–52.

Xu, H., Caimano, M., Lin, T., He, M., Radolf, J., Norris, S., Gherardini, F., Wolfe, A. and Yang, X. (2010) 'Role of acetyl-phosphate in activation of the Rrp2-RpoN-RpoS pathway in Borrelia burgdorferi', *PLoS pathogens.*, 6 (9).

Xu, H., He, M., He, J. and Yang, X. (2010) 'Role of the surface lipoprotein BBA07 in the enzootic cycle of Borrelia burgdorferi', *Infection and immunity.*, 78 (7): 2910–8.

Yamakura, F., Kobayashi, K., Furukawa, S. and Suzuki, Y. (2007) 'In vitro preparation of iron-substituted human manganese superoxide dismutase: Possible toxic properties for mitochondria', *Free radical biology & medicine.*, 43 (3): 423–30.

Yamashita, S., Lukacik, P., Barnard, T., Noinaj, N., Felek, S., Tsang, T., Krukonis, E., Hinnebusch, B. and Buchanan, S. (2011) 'Structural insights into Ail-mediated adhesion in Yersinia pestis', *Structure (London, England : 1993).*, 19 (11): 1672–82.

Yang X., Promnares K., Qin J., He M., Shroder D.Y., Kariu T., Wang Y., Pal U. (2011) 'Characterization of multiprotein complexes of the*Borrelia burgdorferi* outer membrane vesicles', *J. Proteome Res.*10:4556–4566.
Yeowell, H. N., and White, J. R., (1982) 'Iron requirement in the bactericidal mechanism of streptonigrin', *Antimicrob Agents Chemother*, 22(6): 961-968.

Yost F. J., Jr., Fridovich I. (1973) An iron-containing superoxide dismutase from *Escherichia coli*. J. Biol. Chem. 248, 4905–4908.

Youn H. D., Kim E. J., Roe J. H., Hah Y. C., Kang S. O. (1996) A novel nickel-containing superoxide dismutase from Streptomyces spp. Biochem. J. 318, 889–896.

Zambrano, M., Beklemisheva, A., Bryksin, A., Newman, S. and Cabello, F. (2004) 'Borrelia burgdorferi binds to, invades, and colonizes native type I collagen lattices', *Infection and immunity.*, 72 (6): 3138–46.

Zeman, P. and Benes, C. (2013) 'Spatial distribution of a population at risk: An important factor for understanding the recent rise in tick-borne diseases (Lyme borreliosis and tick-borne encephalitis in the Czech Republic)', *Ticks and tick-borne diseases.*, 4 (6): 522–30.

Zeth, K., Diederichs, K., Welte, W. and Engelhardt, H. (2000) 'Crystal structure of Omp32, the anion-selective porin from Comamonas acidovorans, in complex with a periplasmic peptide at 2.1 A resolution',*Structure (London, England : 1993).*, 8(9), pp. 981–92.

Zhang, J. and Norris, S. (1998) 'Kinetics and in vivo induction of genetic variation of vlsE in Borrelia burgdorferi', *Infection and immunity.*, 66(8), pp. 3689–97.

Zhang, L., Zhang, Y., Adusumilli, S., Liu, L., Narasimhan, S., Dai, J., Zhao, Y.O. and Fikrig, E. (2011) 'Molecular interactions that enable movement of the Lyme disease agent from the tick gut into the Hemolymph', *PLoS Pathogens*, 7(6), p. e1002079.

Chapter 6: Appendix

6.1. Appendix Contents

Appendix 1: Sanger Sequencing

Sequencing data and chromatograms for *B. burgdorferi* OmpA-like proteins

Raw FASTA sequencing data for *B. burgdorferi* OmpA-like proteins

Appendix 2: Circular Dichroism

Circular dichroism data for *B. burgdorferi s.l.* OmpA-like proteins

Circular dichroism reference sets

Appendix 3: BLAST searching key *Borrelia* proteins

Protein BLAST results for Borrelia housekeeping proteins

Appendix 4: Publications

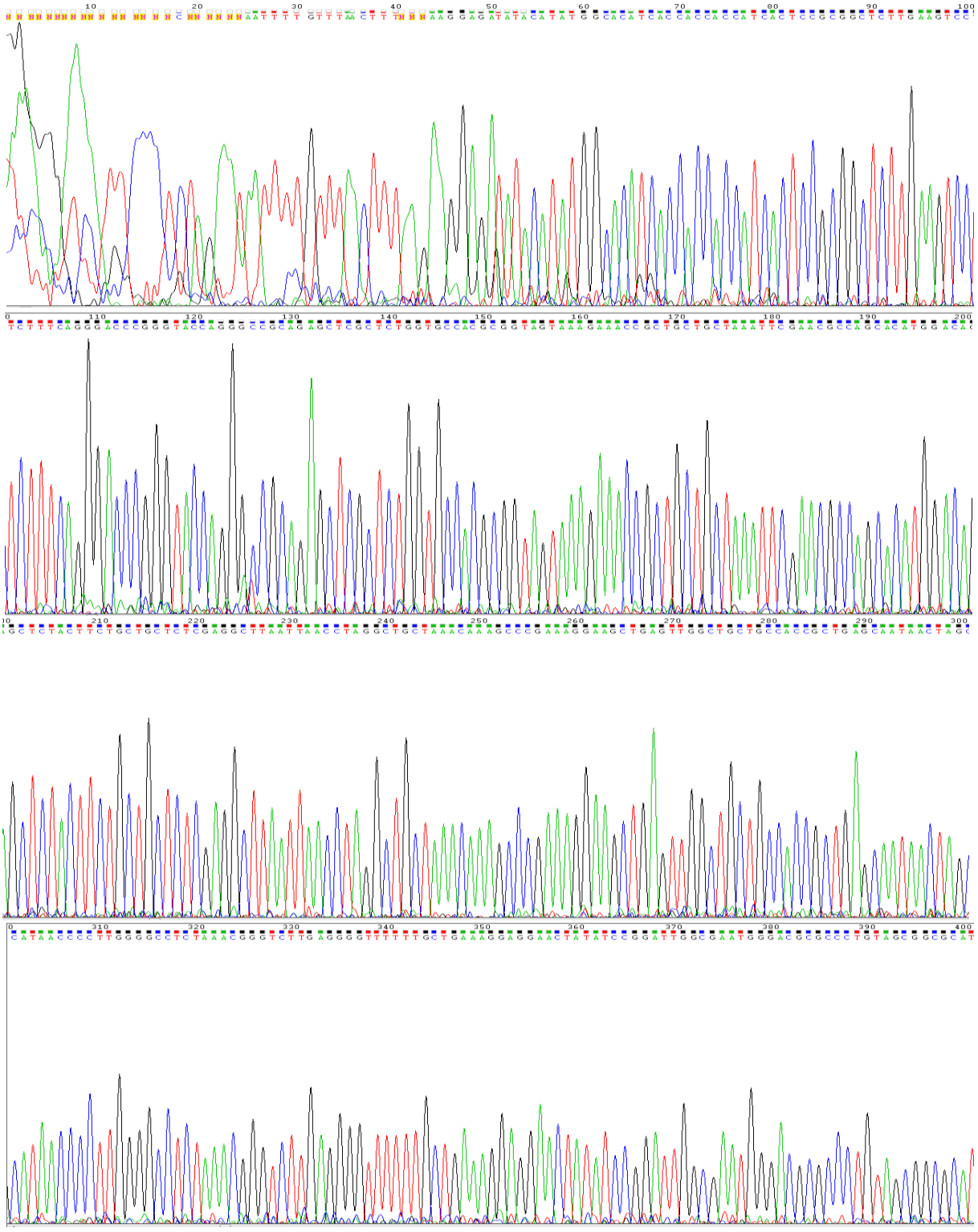
The Borrelia afzelii outer membrane protein BAPKO_0422 binds human factor-H and is predicted to form a membrane-spanning β-barrel.

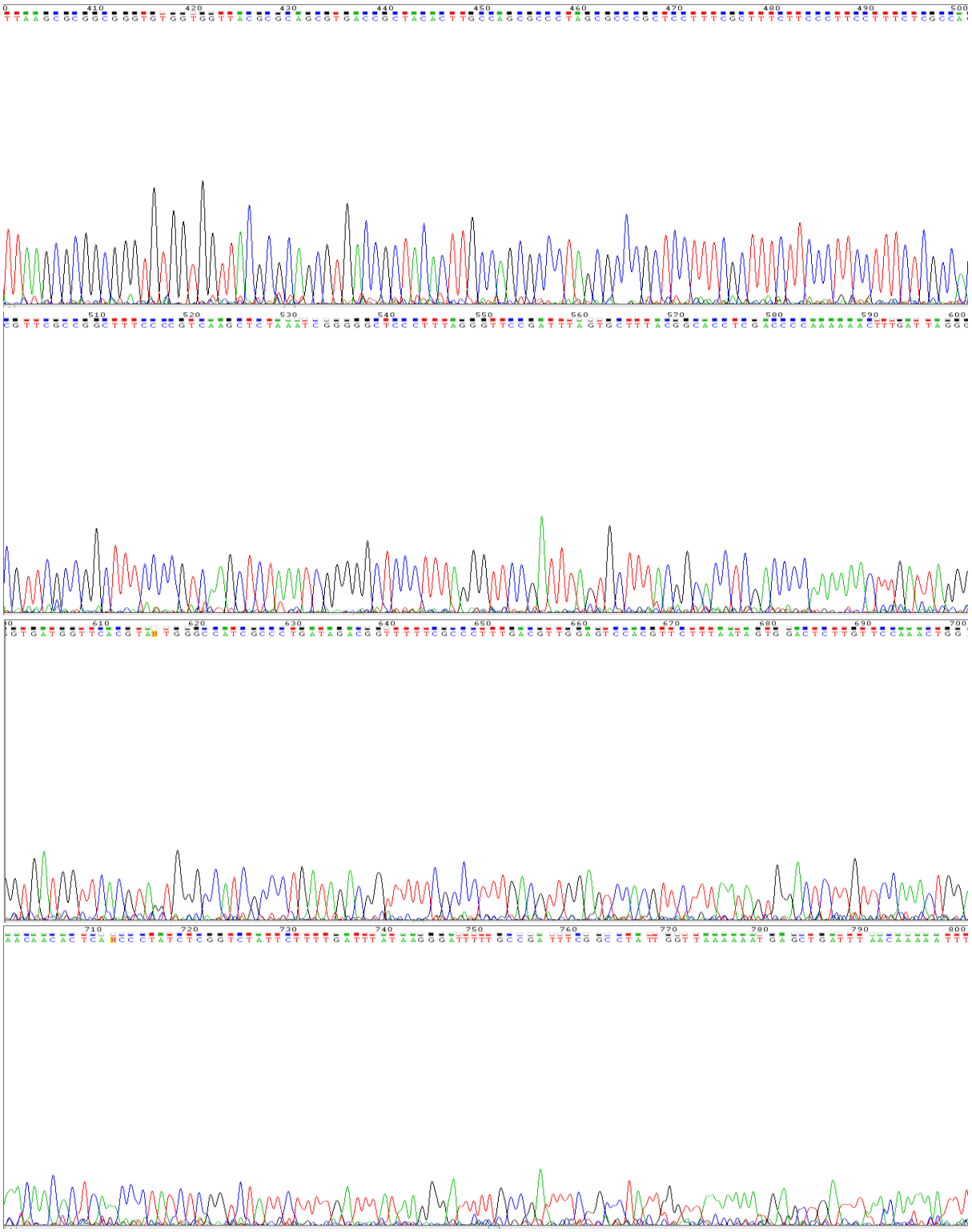
Appendix 1.

Sequencing data and chromatograms for *B. burgdorferi* OmpA-like proteins

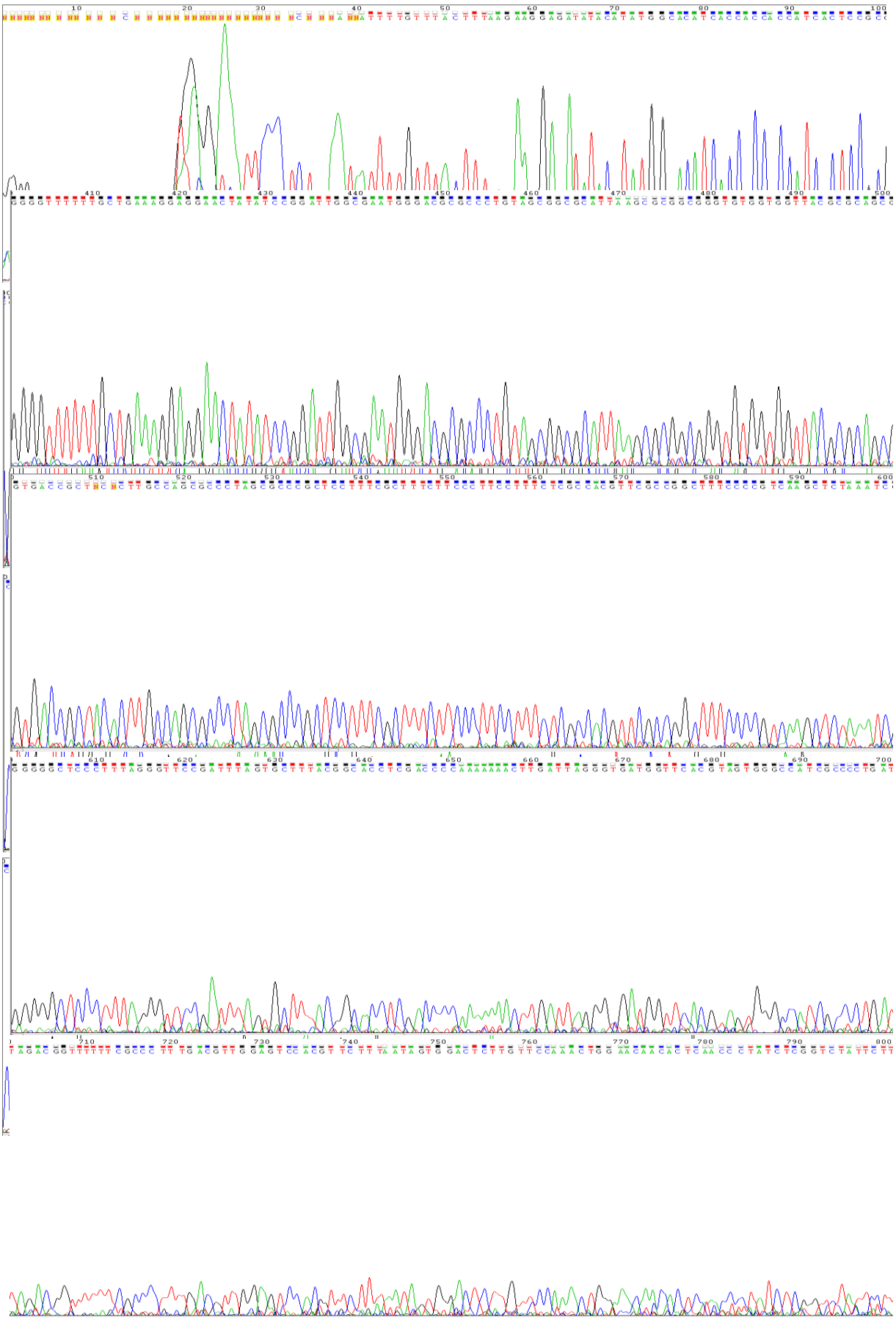
Sequencing was performed with both forward and reverse using primers for the T7 promoter and terminator respectively. All sequencing were performed by Source Bioscience (Rochdale).

Chromatogram 1. BB_0562 forward read

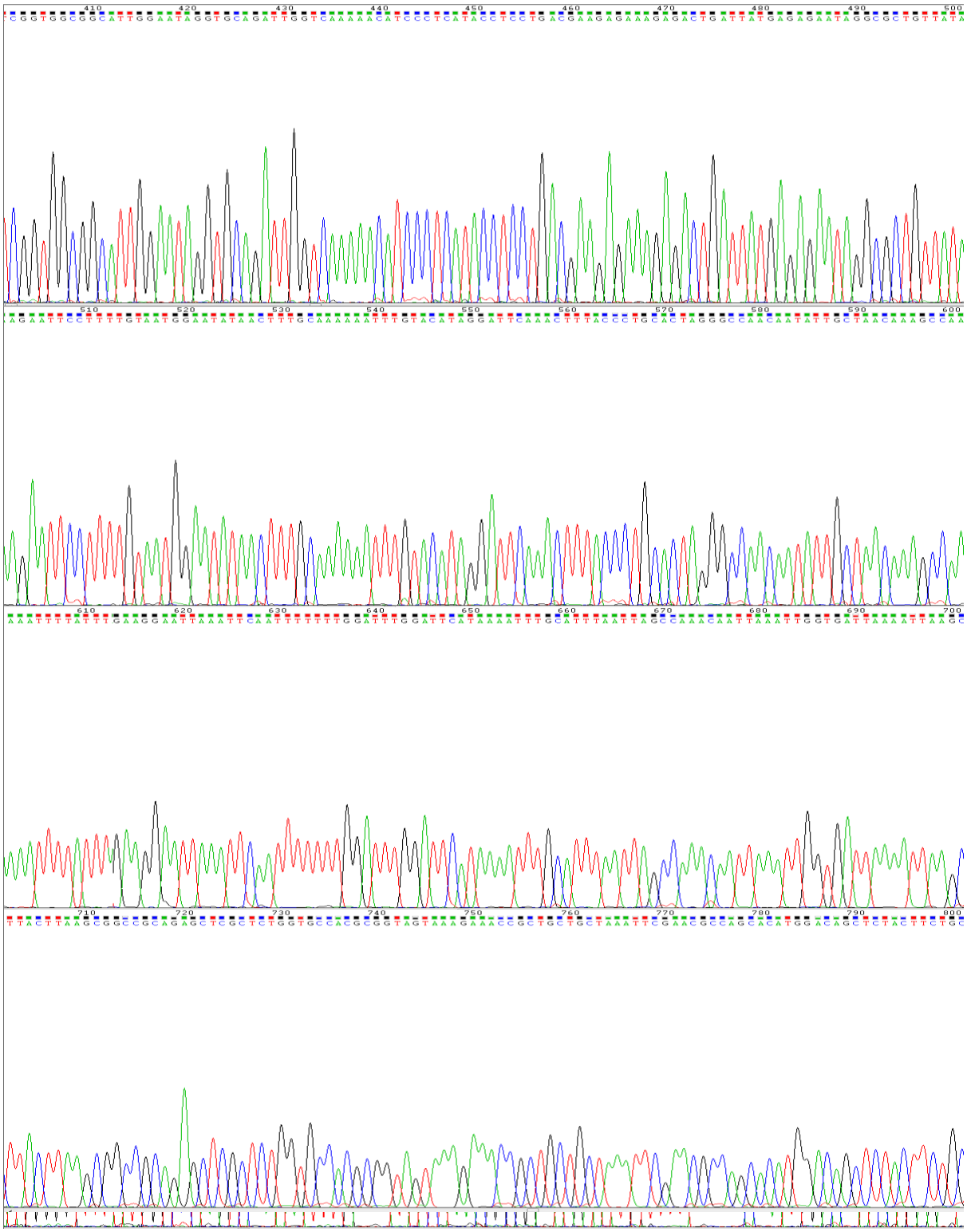




Chromatogram 2. BB_0406 forward read



Chromatogram 3. BG0408 forward read



Raw FASTA sequencing data for *B. burgdorferi* OmpA-like proteins

BB_0562 – Forward

NNNNNNNNNCCCTCTANNNNNTTTTGTTNNCTTTAAGAAGGAGATATACATATGGCACATCACCACCACCATCACTCCGC
GGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCAGGATCCGAAAGATTCATATTTAAATAGAGGAATTGGCTTTGGAG
CAAGCATTGGAAATCCAATTATTAACTTAATAATGTCATTTCCTTTCATTGATTTTGAAATTGGCTATGGTGGTAGTAAT
GGAATAAATCTATCAGGCCCCAAACTTGAATCAAAATTTTATGATTTTAATTTATTAGCAATAGCAGCACTTGATTTCAT
TTTTACAATATCTTTGATAAAAAATTTAAATTTAGGAATTGGAATAGGAGGAAATATAAGCATATCGTCTCACACATCTA
AATTAATAAATGTAGAATTAGGATTTGGAATGAGAATTCCATTGGTTATTTTTTACGACATTACAGAAAATTTAGAAATA
GGTATGAAAATAGCACCTTCAATAGAATTCATCTCAAATACAAGGTCTCTTGCTCAACATAGAACCTATTCGGGCATAAA
ATCAAACTTTGCTGGGGGAATATTTGCTAAGTACTATATCTTTTAACACCAATTCATTCTAATTTATTAAATTTATTGCC
AAGCGCTATGAAAAAATTCGCGGCCGCAGAGCTCGCTCTGGTGCCACGCGGTAGTAAAGAAACCGCTGCTGCTAAATTCG
AACGCCAGCACATGGACAGCTCTACTTCTGCTGCTCTCGAGGCTTAATTAACCTANGCTGCTAAACAAAGCCCGAAAGGA
AGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTT
TGCTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAGCGCGGCGGGNGGNNGGNN
GNTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTANCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCAC
GTTCGCCGGCTTTCCCCGTCNNCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTANNGCTTTACGGCNCNCNACCCC
AAAACTGATTANNNGATGNTNNNCGTANNGGGCNTCNCCCNNNNNNANNGNTTTTNNCCCNTNNNNCNNNGNNTCACNNN
NTTTANNANNNGNANNCTTNNNCNNACNGGANANNNNNNNNCNNNNNNNNNNNNTNNNNNNNNAANGGNNNTNNNCNANT
NNNNNNCNNNNGNNNNAAAAANNNNNNNATTNNNNNNTNNNNCNNATTNNNNNNNNNNNNNNNNNNNNNTTNNNNNNN
NNNCGNN

BB_0562 – Reverse

NNNNNNTCNTTCGGGCTTTGTTTAGCAGCCTAGGTTAATTAAGCCTCGAGAGCAGCAGAAGTAGAGCTGTCCATGTGCTG
GCGTTCGAATTTAGCAGCAGCGGTTTCTTTACTACCGCGTGGCACCAGAGCGAGCTCTGCGGCCGCGAATTTTTTCATAG
CGCTTGGCAATAAATTTAATAAATTAGAATGAATTGGTGTTAAAAGATATAGTACTTAGCAAATATTCCCCCAGCAAAGT
TTGATTTTATGCCCGAATAGGTTCTATGTTGAGCAAGAGACCTTGTATTTGAGATGAATTCTATTGAAGGTGCTATTTTC
ATACCTATTTCTAAATTTTCTGTAATGTCGTAAAAAATAACCAATGGAATTCTCATTCCAAATCCTAATTCTACATTTAT
TAATTTAGATGTGTGAGACGATATGCTTATATTTCCTCCTATTCCAATTCCTAAATTTAAATTTTTTATCAAAGATATTG
TAAAAATGAAATCAAGTGCTGCTATTGCTAATAAATTAAAATCATAAAATTTTGATTCAAGTTTGGGGCCTGATAGATTT
ATTCCATTACTACCACCATAGCCAATTTCAAAATCAATGAAAGGAAATGACATTATTAAGTTAATAATTGGATTTCCAAT
GCTTGCTCCAAAGCCAATTCCTCTATTTAAATATGAATCTTTCGGATCCTGGTACCCGGGTCCCTGAAAGAGGACTTCAA
GAGCCGCGGAGTGATGGTGGNGNGGGNGATGTGCCATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAG
GGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATCGATCTCGATCCTCTA
CGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGG
GAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCNNNCCCGNGGNCGGGGGACTG
TTGGGCGCCATCTNCNTGCATGCACCATTCCTTGNNNNGNNGNGCTNANNNNNNNNACTACTGGCTGNTNCTNNNGCAGN
NNNCATAGNANNNCGNNNNATCCCGNNNCCNNNNNNGGNNNNANNTTNCGNNGNATGNTNNNNNNNNNNNNNNANTCNGG
NGNGNANNNNNNNNNACNNNNNANNNNCNAANNNNNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTNNNNNNNNNNNNN
NNNNNNNNNNNNNNTNNTCN

BB_0406 – Forward

NNNNNNNNNNTCCCTCTNNNNNNNTTTTGTTTAACTTTAAGAAGGAGATATACATATGGCACATCACCACCACCATCACT
CCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCAGGATCCGTCTGACAATTATATGGTCAGATGCAGCAAGGAA
GAAGATTCAACCACCTGTATCGCAAAGCTTAAAGAAATAAAAGAAAAGAAAGATTATGACTTATTTTCAATGGGCATTGG
AATAGGAGATCCTATTGCAAATATTATGATTACAATTCCTTATATAAATATTGATTTTGGATGTGGAGGTTTTATTGGCC
TTAAGTCAAACAATTTTGAAAATTATCTAAATGGTGGAATAGACGTTATTTTTAAAAAGCAAATTGGACAATATATGAAA
ATTGGCGGCGGCATTGGAATAGGTGCGGATTGGTCAAAAACATCCCTTATACCCCCTAATGAAGAAGAAGAAACTGATTA
TGAGAGAATAGGCGCTGTTATAAGAATTCCTTTTATAATGGAATATAATTTTGCAAAAAATTTATCCATAGGATTCAAAA
TTTATCCTGCAGTAGGGCCAACAATATTACTAACAAAACCCAGCATTTTATTTGAAGGAATTAAATTCAATTTTTTTGGA
TTTGGATTCATAAAATTTGCATTTAATTAACCAAATAATTAAATTGGTAATTAAAATTAAGCTCGCATAGCGGCCGCAGA
GCTCGCTCTGGTGCCACGCGGTAGTAAAGAAACCGCTGCTGCTAAATTCGAACGCCAGCACATGGACAGCTCTACTTCTG
CTGCTCTCGAGGCTTAATTAACCTAGGCTGCTAAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCA
ATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGG
CGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCA
GCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAAT
CGGGGGCTCCCTTTAGGGTTCCGATTTNNTGCTTTANNNNCNCGACCCCAAAAACTTGATTAGGGTGATGGNNACGTANN
GGGNNTCNCCNGATNNANGNTTTCNCCTNNACNNNGNANNCCACNTNCNTTANANNGNNTNTNTCNACTNNNNANNNNNN
NNNATCTCNGNNNNNTNTTGNNTNNNANGNNTTNNNGANTNNNNNNNNNNNNNAAANNNNNNNNNNNNNNNNNNNNNC
NANTTTNNNNNNN

BB_0406 – Reverse

NNNNNNTCNTTCGGGCTTTGTTTAGCAGCCTAGGTTAATTAAGCCTCGAGAGCAGCAGAAGTAGAGCTGTCCATGTGCTG
GCGTTCGAATTTAGCAGCAGCGGTTTCTTTACTACCGCGTGGCACCAGAGCGAGCTCTGCGGCCGCTATGCGAGCTTAAT
TTTAATTACCAATTTAATTATTTGGTTAATTAAATGCAAATTTTATGAATCCAAATCCAAAAAAATTGAATTTAATTCCT
TCAAATAAAATGCTGGGTTTTGTTAGTAATATTGTTGGCCCTACTGCAGGATAAATTTTGAATCCTATGGATAAATTTTT
TGCAAAATTATATTCCATTATAAAAGGAATTCTTATAACAGCGCCTATTCTCTCATAATCAGTTTCTTCTTCTTCATTAG
GGGGTATAAGGGATGTTTTTGACCAATCCGCACCTATTCCAATGCCGCCGCCAATTTTCATATATTGTCCAATTTGCTTT
TTAAAAATAACGTCTATTCCACCATTTAGATAATTTTCAAAATTGTTTGACTTAAGGCCAATAAAACCTCCACATCCAAA
ATCAATATTTATATAAGGAATTGTAATCATAATATTTGCAATAGGATCTCCTATTCCAATGCCCATTGAAAATAAGTCAT
AATCTTTCTTTTCTTTTATTTCTTTAAGCTTTGCGATACAGGTGGTTGAATCTTCTTCCTTGCTGCATCTGACCATATAA
TTGTCAGACGGATCCTGGTACCCGGGTCCCTGAAAGAGGACTTCAAGAGCCGCGGAGTGATGGTGGGNGGTGATGTGCCA
TATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGA
GTCGTATTAATTTCGCGGGATCGAGATCGATCTCGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACA

GGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTG
TTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGNNGGGGGANTGTTGGGCGCCATCTNCNTGCATGCANCATTCCTTNNNN
NGNNGGTGNTCANCGGNNTNANCTACTACTGGGCTGCTTCCTAATGCAGNANTCNNATANNANNNCGTCNAGATCCCNGA
NNNCATCNNNGNNNAAACNTNNNNGNNNNGCATGNTANCNCCNNANNNNNTCATNNGGNNGGNNANNNNANNGTANNTNT
NNNANNNNCNNNNNNNCNNNNCTTNNNNNACNNNNCNNNNNGNNNNCNNNNNNNNNNNNNNNNNNNAANNNNNNNTGNGA
NNNNTNNNNNNNNNNNNNNNNNN

BG0408 – Forward

NNNNNNNNNNCNNNCTNNNNNNATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCACATCACCACCACCATCACTC
CGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCAGGATCCGTCTGACAATTATATGGTCAGACGCAGCAAGGAAG
AAGATTCAACAACCTGTATCGCAAAGCTTAAAGGCATAAAAGAAAAAAAAAGTTATGACTTATTTTCAATGGGTATTGGA
ATAGGCAATCCTATTGCAAACATAATAATTACAATTCCTTATATAAATATTGATTTTGGATATGGGGGTTTTATTGGTCC
TAAGTCAAATAATTTTGAAAATTATCTAAACGGCGGAATAGATATTGTTTTTAAAAAACAAATTGGACAATACATGAGAA
TCGGTGGCGGCATTGGAATAGGTGCAGATTGGTCAAAAACATCCCTCATACCTCCTGACGAAGAGAAAGAGACTGATTAT
GAGAGAATAGGCGCTGTTATAAGAATTCCTTTTGTAATGGAATATAACTTTGCAAAAAATTTGTACATAGGATTCAAACT
TTACCCTGCACTAGGGCCAACAATATTGCTAACAAAGCCAAAAATTTTATTTGAAGGAATTAAATTCAATTTTTTTGGAT
TTGGATTCATAAAATTTGCATTTAATTAGCCAAACAATTAAATTGGTGATTAAAATTAAGCTTACTTAAGCGGCCGCAGA
GCTCGCTCTGGTGCCACGCGGTAGTAAAGAAACCGCTGCTGCTAAATTCGAACGCCAGCACATGGACAGCTCTACTTCTG
CTGCTCTCGAGGCTTAATTAACCTANGCTGCTAAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCA
ATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGG
CGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCA
GCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCNGGCTTTCCCCGTCAGCTCTAAATC
GGGGGCTCCCTTTAGGGNTCCGANTTANNGCTTTANGCACNCGACCCCNAAAACTTGATTAGGNTGATNNNCACGTNNNG
NNNTCNNCCNNATANANGTTTTCNCCNTNGACNTNNANTCNNNTNNTTANANNGNNTNNNNNNANCNGANNNNNNNNNNN
TCNNGNNTNNNNTTTGATNNNNNNNNTTTNCNANTTNNNNNNNNTNNANNNNNNTNNNNNNNNNNNNNNNNNNTNANN
NNNTTTNN

BG_0408 – Reverse

NNNNNNNTNNTTCGGGCTTTGTTTAGCAGCCTAGGTTAATTAAGCCTCGAGAGCAGCAGAAGTAGAGCTGTCCATGTGCT
GGCGTTCGAATTTAGCAGCAGCGGTTTCTTTACTACCGCGTGGCACCAGAGCGAGCTCTGCGGCCGCTTAAGTAAGCTTA
ATTTTAATCACCAATTTAATTGTTTGGCTAATTAAATGCAAATTTTATGAATCCAAATCCAAAAAAATTGAATTTAATTC
CTTCAAATAAAATTTTTGGCTTTGTTAGCAATATTGTTGGCCCTAGTGCAGGGTAAAGTTTGAATCCTATGTACAAATTT
TTTGCAAAGTTATATTCCATTACAAAAGGAATTCTTATAACAGCGCCTATTCTCTCATAATCAGTCTCTTTCTCTTCGTC
AGGAGGTATGAGGGATGTTTTTGACCAATCTGCACCTATTCCAATGCCGCCACCGATTCTCATGTATTGTCCAATTTGTT
TTTTAAAAACAATATCTATTCCGCCGTTTAGATAATTTTCAAAATTATTTGACTTAGGACCAATAAAACCCCCATATCCA
AAATCAATATTTATATAAGGAATTGTAATTATTATGTTTGCAATAGGATTGCCTATTCCAATACCCATTGAAAATAAGTC
ATAACTTTTTTTTTCTTTTATGCCTTTAAGCTTTGCGATACAGGTTGTTGAATCTTCTTCCTTGCTGCGTCTGACCATAT
AATTGTCAGACGGATCCTGGTACCCGGGTCCCTGAAAGAGGACTTCAAGAGCCGCGGAGTGATGGTGGTGGTGATGTGCC
ATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTG
AGTCGTATTAATTTCGCGGGATCGAGATCGATCTCGATCCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCAC
AGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTT
GTTTCGGCGTGGGTATGGTGGCAGGCCCCGNGGNNGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCG
CGGCGNGCTCANGGNNCAACNACTACTGNNTGNNNNAANGCNGNNTCGCATNNNANANCGNCNANATCCCGNNNCATCNA
TGNNCAAACNTTNNNNNANGNATGATAGCGCCNNNNNNNANTNNNNNGNNNGNNANNNNAAANNGTANNTNNNNCNANNN
NNCANANNNNCNNNNNNNNNNNTANNNNANNNTTNCCNNNNNNGNNNNNNNNNNNCNNCNNTTCTGNNANNNCNGGAAA
NNNGANNNGNNNNNNNGNNNGNNNNNNNNA

Appendix 2.

Circular dichroism data for *B. burgdorferi s.l.* OmpA-like proteins

CD data were generated for *Borrelia* OMPs across several concentrations the resulting spectra are shown below.

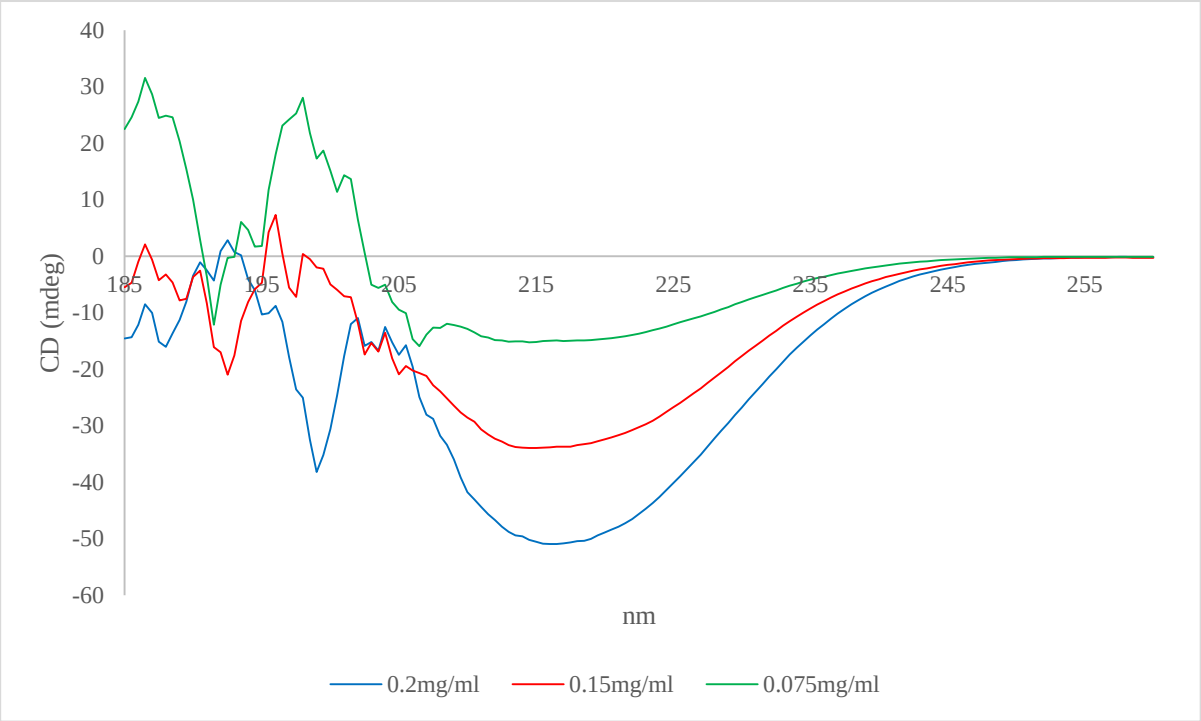


Figure 7.1. CD spectra of BAPKO_0422 at varying concentrations. CD data was collected on a JASCO J-810 at 20ºC from 260 nm down to 185 nm. Recordings were made at 0.5 nm intervals. Proteins were prepared in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO at pH8.

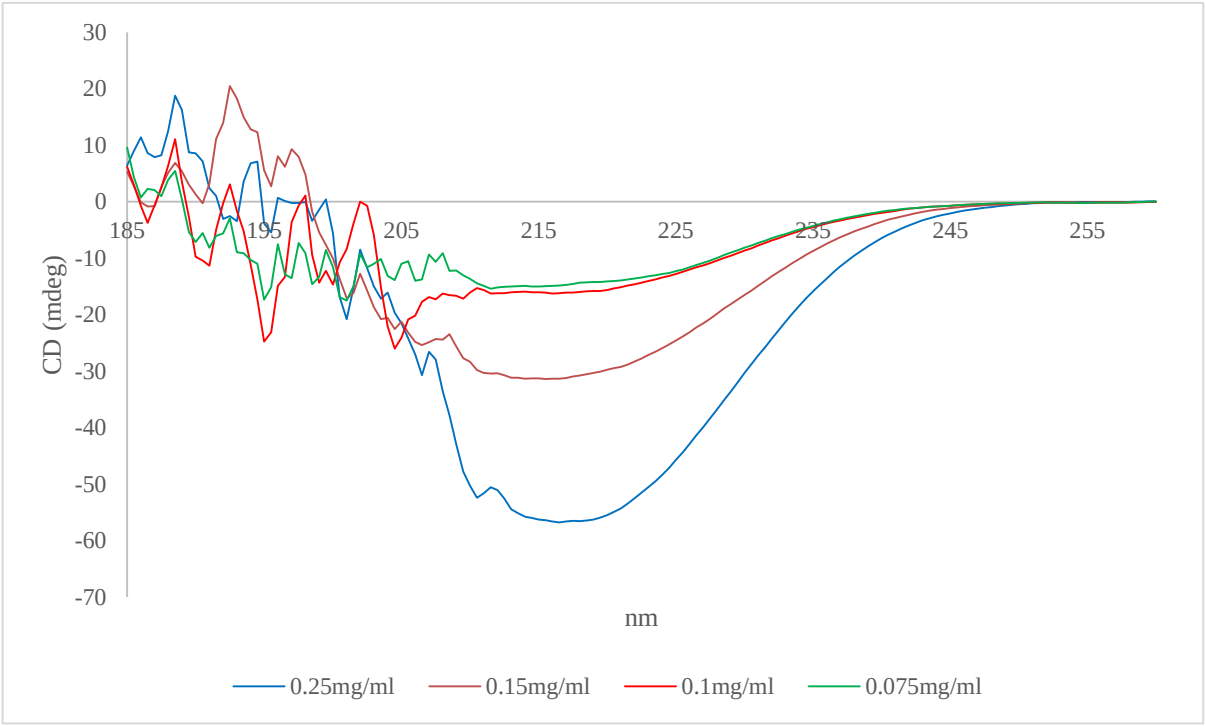


Figure 7.2. CD spectra of BB_0562 at varying concentrations. CD data was collected on a JASCO J-810 at 20ºC from 260 nm down to 185 nm. Recordings were made at 0.5 nm intervals. Proteins were prepared in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO at pH8

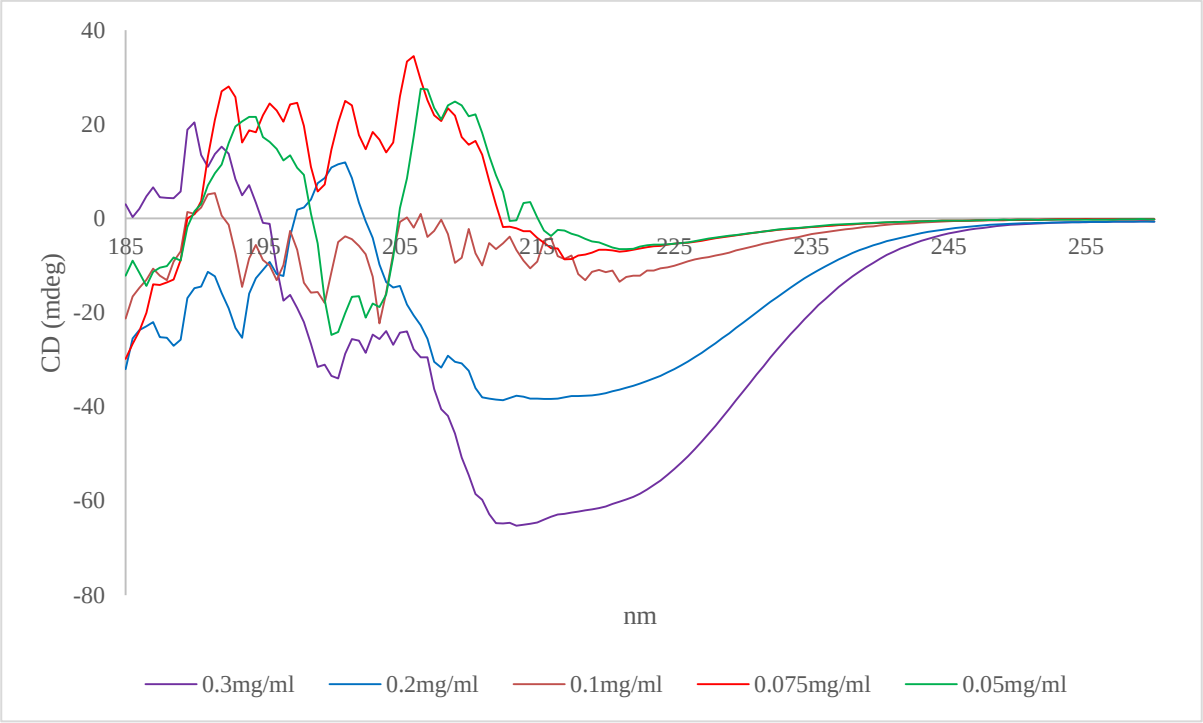


Figure 7.3. CD spectra of BB_0406 at varying concentrations. CD data was collected on a JASCO J-810 at 20ºC from 260 nm down to 185 nm. Recordings were made at 0.5 nm intervals. Proteins were prepared in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO at pH8

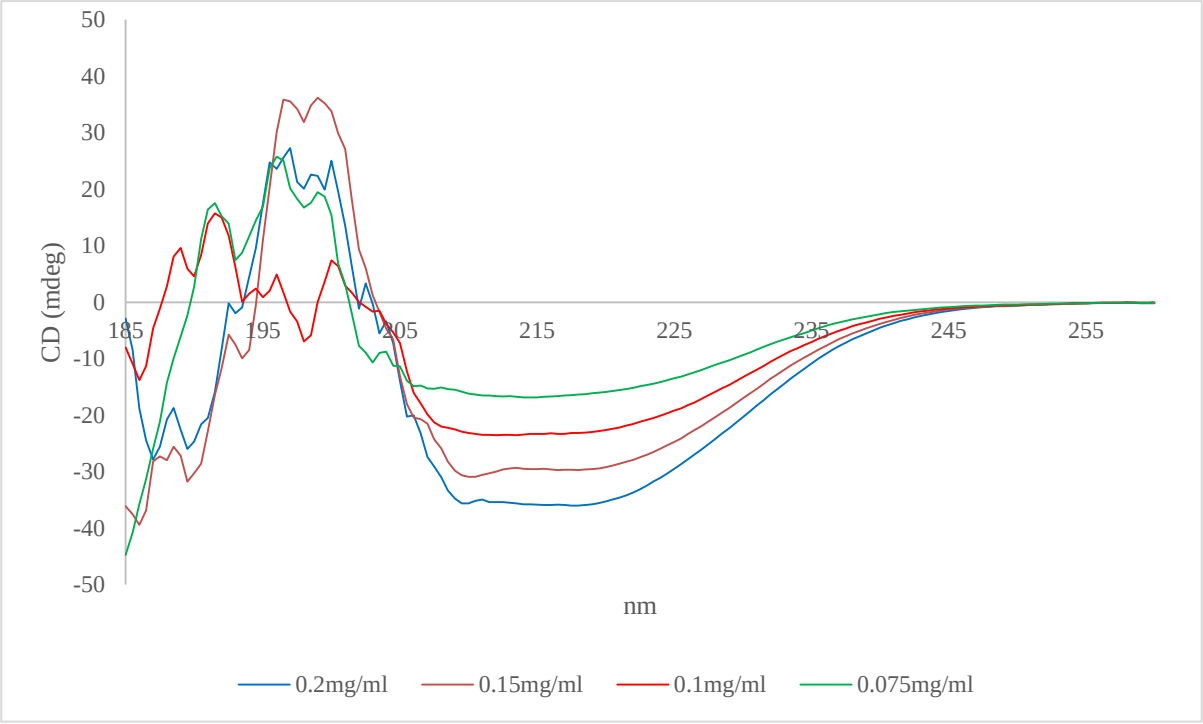


Figure 7.4. CD spectra of BG0408 at varying concentrations. CD data was collected on a JASCO J-810 at 20ºC from 260 nm down to 185 nm. Recordings were made at 0.5 nm intervals. Proteins were prepared in 0.3M NaCl, 50mM tris base, 0.1% v/v LDAO at pH8

Circular dichroism reference sets

Reference Set: SP175

Protein	Organism	PDB
Aldolase	*Rabbit*	1ADO
Alkaline phosphatase	*E. coli*	1 E D 9
α-Amylase	*B. licheniformis*	1VJS
α-bungarotoxin	*Many-banded krait*	1HC9
α-chymotrypsin	*Cow*	5CHA
α-chymotrypsinogen	*Cow*	2CGA
Aprotinin	*Cow*	5PTI
Avidin	*Chicken*	1RAV
β-amylase	*Sweet potato*	1FA2
βB2-Crystalin	*Cow*	2BB2
β-galactosidase	*E. coli*	1BGL
β-lactoglobulin	*Cow*	1B8E
Carbonic anhydrase-I	*Human*	1HCB
Carbonic anhydrase-II	*Cow*	1CA2
Carboxypeptidase-A	*Cow*	5CPA
Calmodulin	*Cow*	1LIN
Catalase	*Human*	1DFG
Ceruloplasmin	*Human*	1KEW
Citrate synthase	*Pig*	2CTS
Concanavalin A	*Jack bean*	1NLS
c-Phycocyanin	*Spirulina platensis*	1HA7
Cytochrome-c	*Horse*	1HRC
Dehydroquinase-typeI	*S. typhimurium*	1QFE
DHQase-II	*M. tuberculosis*	2DHQ
DNase-I	*Cow*	3DNI
Elastase	*Pig*	3EST
Ferredoxin	*C. pasteurinum*	1FDN
γ-B-Crystalin	*Cow*	4GCR
γ-D-Crystalin	*Cow*	1ELP
γ-D-Crystalin	*Human*	1HK0
γ-E-Crystalin	*Cow*	1M8U
γ-S-Crystalin (C-term)	*Human*	1HA4
Glucose oxidase	*Aspergillus niger*	1CF3
Glutamate dehydrogenase	*Cow*	1HWX
Glycogen phosphorylase-b	*Rabbit*	1GPB
Haloalkane dehydrogenase	*R. spheroids*	1BN6
Hemoglobin	*Cow*	1HDA
IgG	*Mouse*	1IGT
Insulin	*Human*	1TRZ
Jacalin	*Jack fruit*	1KU8
Kap95p	*Baker's yeast*	2BKU
Lactoferrin	*Cow*	1BLF
Latexin	*Mouse*	1WNH
Lentil lectin	*Common lentil*	1LES
Leptin	*Human*	1AX8
Lysozyme	*Chicken*	1 9 3 L
Monelin	*Serendipity berry*	1M OL
Myoglobin	*Horse*	1YMB
Myoglobin	*Sperm whale*	1 A 6 M
NmrA	*A. nidulans*	1 K 6 J
Ovalbumin	*Chicken*	1OVA
Ovotransferrin	*Duck*	1DOT
Papain	*Papaya*	1PPN
Pea lectin	*Garden pea*	1OFS

Pepsinogen	*Pig*	2PSG
Peroxidase	*Horseradish*	7ATJ
Phosphoglucomutase	*Rabbit*	3PMG
Phoshoglycerate kinase	*Baker's yeast*	3PGK
Phospholipase-A2	*Cow*	1UNE
PNMT	*Human*	1HMN
Pyruvate kinase	*Rabbit*	1A49
Ribonuclease A	*Cow*	3RN3
Rubredoxin	*Desulfovibrio vulgaris*	1RB9
Serum Albumin	*Human*	1N5U
Streptavidin	*Strep avidinii*	1STP
Subtilisin A	*B. licheniformis*	1SCA
Superoxide dismutase (Cu, Zn)	*Cow*	1CBJ
Thaumatin	*Thaumatococcus daniellii*	1THW
Triose phosphate isomerase	*Baker's yeast*	7TIM
Trypsin inhibitor	*Soybean*	1AVU
Ubiquitin	*Human*	1UBI

Reference Set: SMP180
All of the proteins in SP175 plus,

Protein	**Organism**	**PDB**
Bacteriorhodopsin	*Halobacterium salinarum*	1QHJ
E203Q mutant of Cl-/H+ exchanger	*E. coli*	2FED
Lactose permease	*E. coli*	2CFQ
C123aT mutant of Succinyl-CoA Synthetase	*E. coli*	2NU9
Ammonia channel	*E. coli*	2NOP
Rhodopsin	*Bos taurus*	1HZX
ABC transporter	*E. coli*	1L7V
SecY	*Methanococcus jannaschii*	1RH5
Cytochrome C oxidase	*Bos taurus*	2DYR
Potassium channel	*Streptomyces lividans*	1J95
MscL	*Mycobacterium tuberculosis*	2OAR
AcrB	*E. coli*	2GIF
Photosynthetic reaction center	*Rhodobacter sphaeroides*	1PCR
Photosynthetic reaction center	*Blastochloris viridis*	2WJN
Cytochrome BC1 complex	*Bos taurus*	1BE3
Ca2+-ATPase Ca2-E1-AMPPCP form	*Oryctolagus cuniculus*	1T5S
Succinate dehydrogenase	*E. coli*	1NEK
Potassium channel KirBac3.1	*Magnetospirillum magnetotacticum*	1XL4
WZA membrane protein	*E. coli*	2J58
Outer membrane complex of type IV secretion system	*E. coli*	3JQO
Translocator domain of NALP autotransporter	*Neisseria meningitidis*	1UYN
Argininosuccinate lyase	*Homo sapiens*	1AOS
Ferric hydroxamate uptake receptor	*E. coli*	2FCP
Ferric enterobactin receptor	*E. coli*	1FEP
Cobalamin transporter	*E. coli*	1NQH
OPCA adhesion	*Neisseria meningitidis*	2VDF
YajI lipoprotein	*E. coli*	2JWY
Rhodopsin II-transducer complex	*Natronomonas pharaonis*	1H2S
Na+/Cl- dependent neurotransmitter transporter	*Aquifex aeolicus*	2A65
LH2 B800-850 Membrane protein	*Rhodoblastus acidophilus*	1NKZ

Appendix 3: BLAST searching key *Borrelia* proteins

Protein name	NCBI accession	Closest Match (organism)	Identity
Aminopeptidase	YP_008686584.1	*Psychromonas sp.*	44%
L-Lactate dehydrogenase	WP_010889680	*Thalassobacillus devorans*	52%
Phosphoglycerate kinase	WP_002658313	*Spirochaeta thermophila*	62%
Ribonuclease III	WP_050581418	*Salinispira pacifica*	49%
UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	AIU79269	*Spirochaeta alkalica*	39%
DNA polymerase I	WP_002557136	*Spirochaeta thermophile*	40%
Uridine kinase	CAA66081	*Salinispira pacifica*	60%
NAD kinase	WP_002660542	*Clostridium sp.*	33%
Glycerol kinase	WP_002663811	*Leptolyngbya sp.*	63%
Acetate kinase	EEH00007	*Spirochaeta cellobiosiphila*	48%
Pyridoxal kinase	AIU79421	*Peptostreptococcus anaerobius*	36%
Telomere resolvase	AIU80041	*Agrobacterium tumefaciens*	27%
Caax protease	WP_002557180	*Spirochaeta thermophila*	39%
Clp protease	WP_044273071	*Sulfurospirillum sp.*	32%
Lon protease	AAB72011	*Spirochaeta bajacaliforniensis*	52%
Zinc protease	WP_010889761	*Filimonas lacunae*	30%
Membrane protease	AIU78903	*Spirochaeta thermophile*	42%
ATP-dependent protease	ADQ29063	*Thiomicrospira sp.*	60%
Heat shock protein	AAB04621	*Thiomicrospira sp.*	60%
Cell division protease	EEF82543	*Spirochaeta alkalica*	58%
Molecular chaperone	WP_002557236	*Spirochaeta cellobiosiphila*	70%
Nucleoside transport system permease	AGS66335	*Deferribacter desulfuricans*	29%
Oligopeptide transport system permease	AGS66350	*Spirochaeta thermophila*	53%
Mg^{2+} transport protein	NP_212514	*Fervidicella metallireducens*	48%
Chromate transport protein	AGS66461	*Atopobacter phocae*	40%
Sensor histidine kinase	WP_002661705	*Methyloglobulus morosus*	33%
dTMP kinase	WP_002664672	*E. coli*	92%
Type III pantothenate kinase	WP_044273090	*Dethiosulfatarculus sandiegensis*	35%

UMP kinase	WP_002661420	*Spirochaeta alkalica*	42%
Phosphoglycerate kinase	AAB53931	*Spirochaeta thermophila*	61%
Cytidylate kinase 2	WP_010889829	*Spirochaete thermophila*	47%
Dephospho-CoA kinase	AIU79220	*Thermotoga sp.*	37%
Adenylate kinase	AIU79103	*Thermovirga lienii*	49%
Mevalonate kinase	WP_038378800	*Waddila chondrophila*	33%
FAD-dependent thymidylate synthase	AIU79595	*Treponema denticola*	48%
Hydroxymethylglutaryl-CoA synthase	WP_031542070	*Spirochaeta sp.*	44%
Methylglyoxal synthase	P_002657078	*Psychrilyobacter atlanticus*	67%

Appendix 4: Publications

Biosci. Rep. (2015) / **35** / art:e00240 / doi 10.1042/BSR20150095

Original Paper

OPEN ACCESS

The *Borrelia afzelii* outer membrane protein BAPKO_0422 binds human factor-H and is predicted to form a membrane-spanning β-barrel

Adam Dyer*[1], Gemma Brown*[1], Lenka Stejskal*, Peter R. Laity*† and Richard J. Bingham*[2]

*Department of Biological Sciences, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, U.K.
†Present Address: Department of Materials Science and Engineering, Sir Robert Hadfield Building, Mappin Street, University of Sheffield, Sheffield S1 3JD, U.K.

Synopsis

The deep evolutionary history of the *Spirochetes* places their branch point early in the evolution of the diderms, before the divergence of the present day *Proteobacteria*. As a *spirochete*, the morphology of the *Borrelia* cell envelope shares characteristics of both Gram-positive and Gram-negative bacteria. A thin layer of peptidoglycan, tightly associated with the cytoplasmic membrane, is surrounded by a more labile outer membrane (OM). This OM is rich in lipoproteins but with few known integral membrane proteins. The outer membrane protein A (OmpA) domain is an eight-stranded membrane-spanning β-barrel, highly conserved among the *Proteobacteria* but so far unknown in the *Spirochetes*. In the present work, we describe the identification of four novel OmpA-like β-barrels from *Borrelia afzelii*, the most common cause of erythema migrans (EM) rash in Europe. Structural characterization of one these proteins (BAPKO_0422) by SAXS and CD indicate a compact globular structure rich in β-strand consistent with a monomeric β-barrel. *Ab initio* molecular envelopes calculated from the scattering profile are consistent with homology models and demonstrate that BAPKO_0422 adopts a peanut shape with dimensions 25×45 Å (1 Å = 0.1 nm). Deviations from the standard C-terminal signature sequence are apparent; in particular the C-terminal phenylalanine residue commonly found in *Proteobacterial* OM proteins is replaced by isoleucine/leucine or asparagine. BAPKO_0422 is demonstrated to bind human factor H (fH) and therefore may contribute to immune evasion by inhibition of the complement response. Encoded by chromosomal genes, these proteins are highly conserved between *Borrelia* subspecies and may be of diagnostic or therapeutic value.

Key words: β-barrel, *Borrelia*, Lyme disease, outer membrane protein A (OmpA), small angle X-ray scattering (SAXS), spirochete.

Cite this article as: Bioscience Reports (2015) **35**, e00240, doi:10.1042/BSR20150095

INTRODUCTION

Various species in the genus *Borrelia* are capable of zoonotic infection in humans when transmitted via the saliva of haematophagous ticks resulting in Lyme disease [1]). Considerable research has been directed at *Borrelia burgdorferi sensu stricto,* the most prevalent strain in North America but only a minor contributor to incidence rates in Europe and Asia. The available data on the natural hosts/vectors present in Europe show that the most common species present are *Borrelia garinii* and *Borrelia afzelii*, the latter of which is responsible for the vast majority of observed erythema migrans (EM) rash [2,3]. Although early treatment with antibiotics can be effective, symptoms may continue post-treatment probably due to a variety factors including the presence of immunogenic cell debris, survival of viable spirochetes [4–6] or the presence of persister phenotypes [7]. *Borrelia* have evolved numerous strategies to aid survival for extended periods within the mammalian host. Evasion strategies include antigenic variation [8,9], variable expression of surface antigens [10–12], invasion of immune privileged sites [13,14] and specific binding of immune regulators [15]. As part of the latter, the

Abbreviations: ALBI, affinity ligand-binding immunoblot; AtVDAC1, *Arabidopsis thaliana* voltage-dependent anion channel 1; BamA, barrel assembly machinery; BB, *Borrelia burgdorferi*; BG, *Borrelia garinii*; BAPKO, *Borrelia afzelii* strain PKO; DDAO, *N*,*N*-dimethyldodecylamine-*N*-oxide; DipA, dicarboxylate-specific porin A; EM, erythema migrans; fH, factor H; HMM, hidden Markov model; OM, outer membrane; OmpA, outer membrane protein A; R_g, radius of gyration; TM, transmembrane.

[1] These authors contributed equally to the work.

[2] To whom correspondence should be addressed (email r.j.bingham@hud.ac.uk).

ability to bind factor-H (fH) at the bacterial cell surface and so exploit the host's own protection against the complement response is an essential virulence factor to establish infection and is thought to determine host specificity [16]. Adhesion, invasion and immune evasion is mediated primarily by a variety of surface lipoproteins and membrane spanning β-barrels. Other notable surface proteins include the integrin-binding BBB07 [17], the α-helical P13 channel protein [18,19] and the BesABC (*Borrelia* efflux system proteins A, B and C) efflux pump [20,21].

Lipoproteins are readily identified by their characteristic signal sequence, consisting of a positively charged N-region, hydrophobic H-region and a lipobox terminating with a single cysteine amino acid [22]. Genome data [23,24] has shown *Borrelia* has ~105 such proteins, many of which have been studied in detail and have been shown to bind to a wide range of host extracellular matrix proteins [25], immune regulators and cell surface receptors [15].

β-Barrel membrane-spanning proteins pose a significant challenge to prediction methods because they are characterized by short (<10 amino acids) membrane-spanning strands of alternate hydrophobic residues interspersed by highly variable loop regions [26]. The sequence conservation between orthologues can also be very low. Consequently and probably also due to their relative scarcity on the outer membrane (OM), very few β-barrel membrane-spanning proteins have been identified in *Borrelia*. Those that have been identified, BesC (BB0142) [20], dicarboxylate-specific porin A (DipA) (BB0418) [27], P66 (BB0603, Oms66) [28,29] and BamA (barrel assembly machinery; BB0795) [30] are all predicted to form 16–24-stranded porin-type barrels. The smaller barrel-types with 8–10 strands, commonly found in other Gram-negative bacteria have so far remained undetected in *Borrelia*.

Highly conserved among the *Proteobacteria*, the outer membrane protein A (OmpA)-like transmembrane (TM) domain, defined by Pfam family PF01389, consists of eight membrane-spanning antiparallel β-strands with four large extracellular loops and three short periplasmic turns [26]. High-resolution structural data are available for several members of this family with a range of different functions [31–34]. This structural data confirmed that the TM region is highly conserved and functional diversity is mainly due to variations in extracellular loop regions. The prototypical *Escherichia coli* OmpA functions as a non-specific diffusion channel and with the addition of a C-terminal peptidoglycan-binding domain, contributes to the maintenance of cell shape and structural integrity of the OM [35].

In addition to these physiological functions, numerous members of this family are known virulence factors including *E. coli* OmpA [36,37], *Cronobacter sakazakii* OmpA [38], *Salmonella typhimurium* Rck and PagC [39,40] and *Yersinia pestis* Ail [31]. These proteins contribute to pathogenicity by binding to a range of host factors mediating adhesion, invasion and resistance to the complement response.

To date, the OmpA-like TM domain has not been identified in *Borrelia*. Considering the similarities between the membrane structure of *Borrelia* and that of the *Proteobacteria*, we considered it likely that the OmpA-like TM domain would be conserved in both. To this aim we conducted PSI-BLAST searches of the NCBI protein database against the available *B. burgdorferi* proteomes using a variety of OmpA sequences as search targets. Although larger β-barrel proteins such as BamA (BB0795) were found, no proteins matching the criteria of an OmpA-like domain could be identified. Consequently, we used a profile hidden Markov model (HMM) search strategy, which enables sequences with more remote homology to be detected [41]. Four homologous proteins were identified, all encoded by chromosomal genes and highly conserved between different species of *Borrelia*. Analysis of signal sequences and bioinformatic analysis suggest that all four proteins adopt similar topology as eight-stranded membrane-spanning β-barrels. As the most common cause of EM in Europe, we chose the *B. afzelii* homologue BAPKO_0422 for further experimental work.

In the following work, we describe the production of a recombinant version of BAPKO_0422 consisting of the predicted membrane-spanning domain (no signal sequence) in the *E. coli* expression system. Furthermore an affinity ligand-binding immunoblot (ALBI) assay demonstrates that BAPKO_0422 binds specifically to human fH. Finally, data from SAXS, CD and phase partitioning demonstrate that BAPKO_0422 adopts a compact peanut-shaped structure rich in β-strand, consistent with a membrane-spanning β-barrel topology.

MATERIALS AND METHODS

Identification of *Borrelia* OmpA-like domains

Protein sequences of all available OmpA-like TM domains were obtained from the Pfam database family PF01389 [42]. Sequences with greater than 90% or 40% similarity were removed using the Decrease Redundancy program (ExPASy) to generate two lists, which were then submitted to JACOP for classification [43]. Three families were identified corresponding to OmpA, OmpX and OmpW. Sequences corresponding to each family were aligned using ClustalW2 (EBI) and used to build a profile HMM using HMMER (http://hmmer.janelia.org/). This generated six profile HMMs, which were then used to search the genomes of *B. afzelii*, *B. burgdorferi* and *B. garinii* to generate a list of potential OmpA-like proteins. The top scoring hits were filtered based on the presence of a signal sequence predicted using Signal P 4.1 [44] and fold prediction using the Fold and Function Assignment System server (FFAS03) [45]. The procedure identified two proteins BB0027 and BAPKO_0422 matching all criteria. A PSI-BLAST search was then conducted to identify homologous proteins in *B. afzelii*, *B. burgdorferi* and *B. garinii*. Topology predictions were conducted using the PRED-TMBB web server [46] using the mature sequences of all homologues identified above (minus signal peptide).

Homology modelling

Homology models were generated using Modeller [47] based on co-ordinates from PDB accession codes 1BXW, 1P4T, 1QJ8,

1QJP, 1THQ, 2ERV, 2F1T, 2K0L, 2X27, 3DZM, 3GP6 and 3QRA. A structure-based multiple sequence alignment was generated using the MatchMaker function in the UCSF Chimera package [48] with default settings. This was manually adjusted to minimize sequence gaps in loop regions. Sequences for the putative *Borrelia* OmpA-like TM domains were then added to this alignment and manually adjusted to match the topology and β-strand positioning as predicted by PRED-TMBB. Multiple models were then generated using Modeller. Theoretical radius of gyration (R_g) values, based on the centre of mass, were calculated using the UCSF Chimera package [48].

Cloning of BAPKO_0422$^{20\text{-}201}$

The coding region for amino acids 20–201 of BAPKO_0422 (European Nucleotide Archive accession number ABH01676.1) was amplified by PCR using the forward primer 5′-GCATGGATCCGCAATCAAAAAGCAAAACTAT-3′ and reverse primer 5′-ATCGAAGCTTTCATTTATTCTCCATTATATATA-3′ with the addition of BamH1 and HindIII restriction sites (underlined). This was ligated into the pET-47b(+) expression vector (Novagen) and confirmed by DNA sequencing. PCR primers were synthesized by Eurofins MWG Operon.

Recombinant expression and purification of BAPKO_0422$^{20\text{-}201}$

Protein expression in Rosetta^{TM}(DE3)pLysS cells (Merck Millipore) was induced by the addition of 1 mM IPTG to a 3 l of culture of cells with a *D* of 0.6. After 3 h, cells were harvested by centrifugation and lysed on ice by pulsed sonication in 0.3 M NaCl, 50 mM Tris/HCl, pH 8.0. The lysate was incubated with DNase before removal of the soluble fraction by centrifugation (20000 ***g*** for 30 min at 4 °C) and the pellet prepared for solubilization by means of several preparative washes (0.3 M NaCl, 50 mM Tris/HCl, 10 mM EDTA, 1 mM DTT, 0.1 % Triton-X-100, pH 8.0). The washed inclusion body was then solubilized in 8 M urea, 0.3 M NaCl, 50 mM Tris/HCl, pH 8.0, overnight and clarified by centrifugation at 16000 ***g*** for 30 min.

Protein was loaded on to a HisTrap HP 5 ml column (GE Healthcare) in 8 M urea, 50 mM Tris/HCl, 0.3 M NaCl at pH 8. On-column refolding was performed by the use of a linear gradient into 1 M urea, 50 mM Tris base, 0.3 M NaCl, 0.1 % *N*,*N*-dimethyldodecylamine-*N*-oxide (DDAO) at pH 8.0 over 60 ml at a flow rate of 0.5 ml/min. Refolded BAPKO_0422$^{20\text{-}201}$ was then subjected to 10 column volumes of wash buffer, 1 M urea, 50 mM Tris/HCl, 1 M NaCl, 50 mM imidazole, pH 8.0, before being eluted by a linear gradient of 50 mM Tris/HCl, 0.3 M NaCl, 0.3 M imidazole, 0.1 % DDAO at pH 8.0.

The N-terminal 6× His-tag was removed enzymatically using the HRV-3 (human rhinovirus-3) protease (Sigma–Aldrich). Purified BAPKO_0422$^{20\text{-}201}$ was incubated at room temperature for 48 h in buffer 0.3 M NaCl, 0.1 % (w/v) DDAO, 50 mM Tris/HCl, pH 8.0. The reaction mixture then was passed through a HisTrap HP 5-ml column (GE Healthcare) to remove the 6× His-tagged HRV-3C and any uncleaved protein.

Protein was concentrated using a 10 NMWL Amicon Ultra-15 centrifugal filter unit (Merck Millipore), centrifuged at 16000 ***g*** for 20 min at 10 °C and the protein concentration measured by spectroscopy and Bradford analysis.

Size exclusion chromatography of BAPKO_0422$^{20\text{-}201}$

Following initial purification, BAPKO_0422$^{20\text{-}201}$ was subjected to gel filtration and was applied to a Superdex 75 10/300 column at a flow rate of 0.5 ml/min in 0.3 M NaCl, 50 mM Tris/HCl, 0.1 % DDAO at pH 8. Peak fractions containing BAPKO_0422$^{20\text{-}201}$ were pooled and concentrated. BSA, ovalbumin, ribonuclease A and vitamin B12 were prepared in the same buffer and used as size exclusion chromatography (SEC) standards.

Affinity ligand-binding immunoblot assays

ALBI assays were performed to detect specific binding to human fH. BAPKO_0422$^{20\text{-}201}$ (10 μg) was subjected to native-PAGE alongside human fH (HC2130, Hycult Biotech) (0.5 μg) as a positive control. Negative controls consisted of BSA (10 μg, Sigma–Aldrich) and recombinant superoxide dismutase A (SodA; 10 μg) from *B. burgdorferi* produced from the same vector and subjected to the same purification protocol. All incubations were carried out at 4 °C. Proteins were immunoblotted on to PVDF membrane and blocked in 5 % milk powder made up in TBS. Following three 5-min washes with TBS-Tween (0.05 %), the membrane was incubated with human fH at a concentration of 73 μg/ml in TBS for 14 h with gentle agitation. This concentration is approximately 3-fold lower than the mean concentration of fH in human blood (233–269 μg/ml) [49]. The membrane was further washed, as described above, before incubation with the primary antibody for 1 h (1:1000 of mouse monoclonal anti-Human fH, Abcam). Following washing as above, the membrane was then subjected to the secondary antibody for 1 h (1:5000 of Goat Anti-Mouse IgG H&L, Alexa Fluor®680, Abcam). The blot was then visualized on a LI-COR Odyssey infrared imaging device by measuring emission at 700 nm. A negative control blot was carried out alongside the experimental assay. In the latter case, human fH was replaced with TBS for the 14-h incubation.

CD spectroscopy

CD data were collected from a solution of 0.33 mg/ml BAPKO_0422$^{20\text{-}201}$ in 0.1 % (w/v) DDAO, 0.3 M NaCl, 30 mM Tris/HCl, pH 8.0, on a Jasco J-810 Spectropolarimeter. Measurements were taken in duplicate over a wavelength range of 195–260 nm before background subtraction using a dialysis matched buffer. Data were processed using the DichroWeb server using the algorithm CDSSTR and reference sets 4 and 7.

Phase partitioning

Phase partitioning was conducted according to published methods [50]. Control proteins haemoglobin and lysozyme were

purchased from Sigma–Aldrich. *Arabidopsis thaliana* voltage-dependent anion channel 1 (AtVDAC1) was refolded and purified according to published methods [51]. All temperature incubations were for a period of 3 min. Protein solutions were made up to 1 mg/ml in 10 mM Tris/HCl, pH 7.4, 150 mM NaCl, 1 % Triton X-114 at 0 °C. A 200 μl of sucrose cushion was prepared consisting of 6 % (w/v) sucrose, 10 mM Tris/HCl, pH 7.4, 150 mM NaCl and 0.06 % Triton X-114. Fifty microlitres of protein solution was overlaid on the sucrose cushion and incubated at 30 °C before centrifugation at 300 ***g*** for 3 min. The upper aqueous layer (70 μl) was removed and additional Triton X-114 was added to 0.5 % concentration. After incubation at 0 °C and gentle mixing, this aqueous phase was overlaid on the same sucrose cushion, incubated at 30 °C and then centrifuged at 300 ***g*** for 3 min to pellet the detergent phase. The upper aqueous phase was removed and received an additional 2 % Triton X-114. This was then cooled to 0 °C, mixed, then incubated at 30 °C before a final centrifugation at 300 ***g*** for 3 min (pellet discarded). The supernatant (aqueous phase) was then analysed by SDS/PAGE alongside the previous detergent phase.

SAXS

SAXS data were acquired using a Bruker Nanostar from solutions of BAPKO_0422$^{20\text{-}201}$ in 0.1 % (w/v) DDAO, 0.3 M NaCl, 50 mM Tris/HCl, pH 8.0. Protein samples of ∼100 μl were sealed in a 1.5 mm bore quartz glass capillary and the sample chamber was evacuated to minimize background scattering. The sample to detector distance was 106.85 cm. The *s*-axis and beam centre were calibrated using the scattering pattern of silver-behenate salt (*d*-spacing = 5.84 nm). The momentum transfer was defined as $s = 4\pi\sin(\theta)/\lambda$. Although detergent solutions exhibit scattering similar to that of water, proper background subtraction is essential to ensure good quality data. In addition, the presence of protein can sequester detergent from the surrounding solution and alter the monomer/micelle equilibrium. For this reason, samples were dialysed against the blank buffer at 4 °C for 5 days prior to data collection. Each scattering data set consisted of 10 × 2400 s exposures. Scattering data were collected from His-tagged BAPKO_0422$^{20\text{-}201}$ at 3.3, 4.5 and 5.5 mg/ml and from native BAPKO_0422$^{20\text{-}201}$ at 2.5, 3.5 and 4.5 mg/ml. For each sample, the equivalent scattering data collected for dialysis-matched buffer (10 × 2400 s) was subtracted using Primus software. Equivalent scattering data were also acquired for lysozyme using identical buffers (result not shown).

R_g values were evaluated using both the Guinier approximation and also in real space from the entire scattering pattern using the indirect transform program GNOM.

Ab initio molecular envelopes were calculated using the DAMAVER program suite, part of the ATSAS package [52,53]. DAMMIN generated 20 independent models by simulated annealing (P1 symmetry) using a bead model, which were then aligned and outliers removed before averaging (Supplementary Figures S1 and S2, Supplementary Tables S1 and S2). A filtered model was generated in DAMFILT. Figures were generated in PyMOL (The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger).

RESULTS

Identification and topology prediction of OmpA-like TM domains in *Borrelia*

The sequences of known OmpA-like membrane-spanning domains, defined by Pfam family PF01389 [42], show high conservation in only a small number of key positions, specifically the various residues involved in forming the two aromatic-girdles, several glycine residues and residues involved in ion pair interactions. The aliphatic residues orientated towards the membrane-interior are highly variable between different proteins whereas the loop-regions are exceptionally variable. Because of this low sequence conservation, we employed a sensitive HMM-based search strategy to find potential OmpA-like domains in the genus *Borrelia*. The results revealed four chromosomally encoded proteins, consisting of three closely related paralogous sequences, BAPKO_0422, BAPKO_0423, BAPKO_0591 and the more distantly related BAPKO_0026. An analysis of whole genome sequences of the 35 species in the Lyme-borreliosis group and seven species in the relapsing fever group [54] revealed conserved orthologues of these proteins in all known *Borrelia* species. Ordered locus names from the three main strains known to be pathogenic to humans (*B. afzelii, B. burgdorferi* and *B. garinii*) are shown in Table 1. All the sequences identified are currently annotated as putative uncharacterized proteins. Fold recognition using the FFAS03 server [45] suggested that BAPKO_0026 was more closely related to *E. coli* OmpW, whereas the BAPKO_0422 paralogous group (BAPKO_0422, BAPKO_0423, BAPKO_0591) bore similarities to a range of proteins including the *Neisseria meningitidis* protein NspA and the *E. coli* proteins OmpA and OmpW (Table 1). However, an analysis of sequence similarity and turn/loop lengths with comparison to the *E. coli* homologues did not reveal any consistent features, suggesting the *Borrelial* proteins share the same basic topology as OmpA/OmpW/NspA but are not closely related to any of these proteins in particular.

Topology predictions (PRED-TMBB) of the proteins listed in Table 1 match an eight-stranded β-barrel with long extracellular loops and short periplasmic turns (Figures 1 and 2). Both N- and C-termini are predicted to be periplasmic. A consensus topology was generated based on the PRED-TMBB results and manual inspection of conserved of aliphatic residues (shaded grey, Figures 2A and 2B). The topology prediction failed for BB0405 and BG0407 and some manual interpretation was required based on the sequence similarity to other members of the group. Some uncertainty remains about the positioning of the first strand (indicated by dashed lines in Figure 2). The sequence alignment reveals that TM regions are more conserved than loop regions, which vary in sequence and length between the four homologues. Side chains predicted to occupy the barrel interior are in general more conserved than neighbouring residues exposed to the hydrophobic lipid-bilayer, as might be expected considering the constraints imposed by packing inside the barrel.

Signal sequence prediction (Signal P) [44] suggests that nine out of the 12 homologues listed in Table 1 contain a functional signal sequence (Figures 2A and 2B). Multiple positively charged

Table 1 Ordered locus names of potential OmpA-like domains identified by HMM/FFAS03/BLAST searches from *B. afzelii* (BAPKO), *B. burgdorferi* (BB) and *B. garinii* (BG)

The top hit from FFAS03 searches are all eight-stranded membrane-spanning proteins (either OmpA, OmpW or NspA). All sequences are currently annotated as putative uncharacterized proteins. Signal sequence prediction results are based on the no TM neural network. *Present as monomer in OM vesicles [65]. †fH-binding activity [59]. ‡UniProt sequence is 11 amino acids shorter than GenBank sequence, full sequence is shown.

Ordered locus name	GenBank (UniProtKB) Accession codes	No. of amino acids	FFAS03 highest scoring template	Signal sequence prediction (SignalP 4.1)
BAPKO_0026	ABH01291.1 (Q0SPD7)	211	OmpW (2X27)	N
BB_0027	AAC66429.1 (O51058)	212	OmpW (2X27)	Y
BG0027	AAU06886.1 (Q662Y5)	212	OmpW (2X27)	N
BAPKO_0422	ABH01676.1 (Q0SNA2)	201	OmpA (2K0L)	Y
BB_0405*	AAC66795.1 (O51366)	203	OmpW (2X27)	Y
BG0407†	AAU07257.1 (Q661L4)	203	OmpW (2X27)	Y
BAPKO_0423	ABH01677.1 (Q0SNA1)	203	OmpA (2K0L)	Y
BB_0406	AAC66794.1 (O51367)	203	OmpA (2K0L)	Y
BG0408	AAU07258.1 (Q661L3)	203	OmpA (2K0L)	Y
BAPKO_0591	ABH01831.1 (G0IQB2‡)	181	NspA (1P4T)	Y
BB_0562	AAC66924.1 (O51510)	180	NspA (1P4T)	N
BG0572	AAU07409.1 (Q660W2)	179	NspA (1P4T)	Y

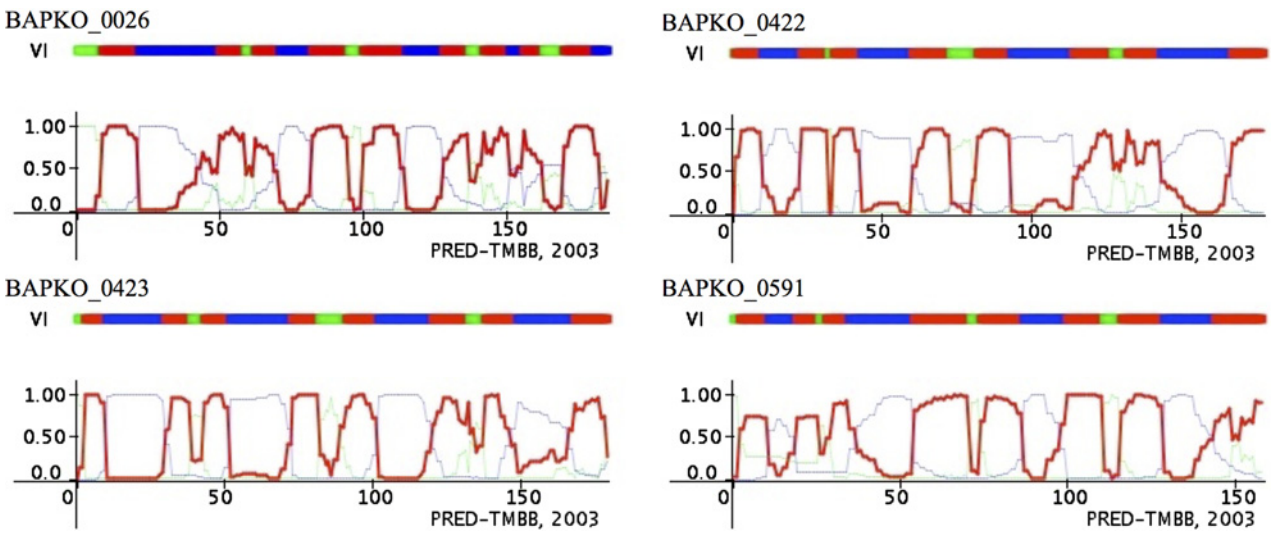


Figure 1 Topology prediction (PRED-TMBB) of ordered loci BAPKO_0026, BAPKO_0422, BAPKO_0423 and BAPKO_0591

Colours indicate regions predicted to be extracellular (blue), periplasmic (green) and TM (red). Sequences are numbered based on the mature protein (no signal sequence).

amino acids are seen in the N-terminal region, a feature consistently found in confirmed signal sequences from both spirochetes and proteobacteria [55,56]. With the exception of BB0562, all sequences have an alanine residue in the predicted −1 position and so are expected to be processed by a type-I signal peptidase.

Inspection of BAPKO_0423, BB0406 and BG0408 shows two cysteine amino acids at positions 31 and 40, indicating potential sites for N-terminal lipidation. However, the potential lipobox is a poor match with previous analysis of spirochetal lipobox sequences [22]. The topology prediction presented in the present study suggest these two cysteine amino acids form part of a small periplasmic region and are therefore likely to be disulfide bonded.

Borrelia OmpA-like domains contain a C-terminal signature sequence

The vast majority of OM β-barrels in Gram-negative bacteria contain a C-terminal signature sequence resembling Z-x-Z-x-Z-x-Y-x-F>, where x represents any amino acid and Z-represents non-polar residues. The terminal residue is invariably a non-polar residue, usually a phenylalanine [57].

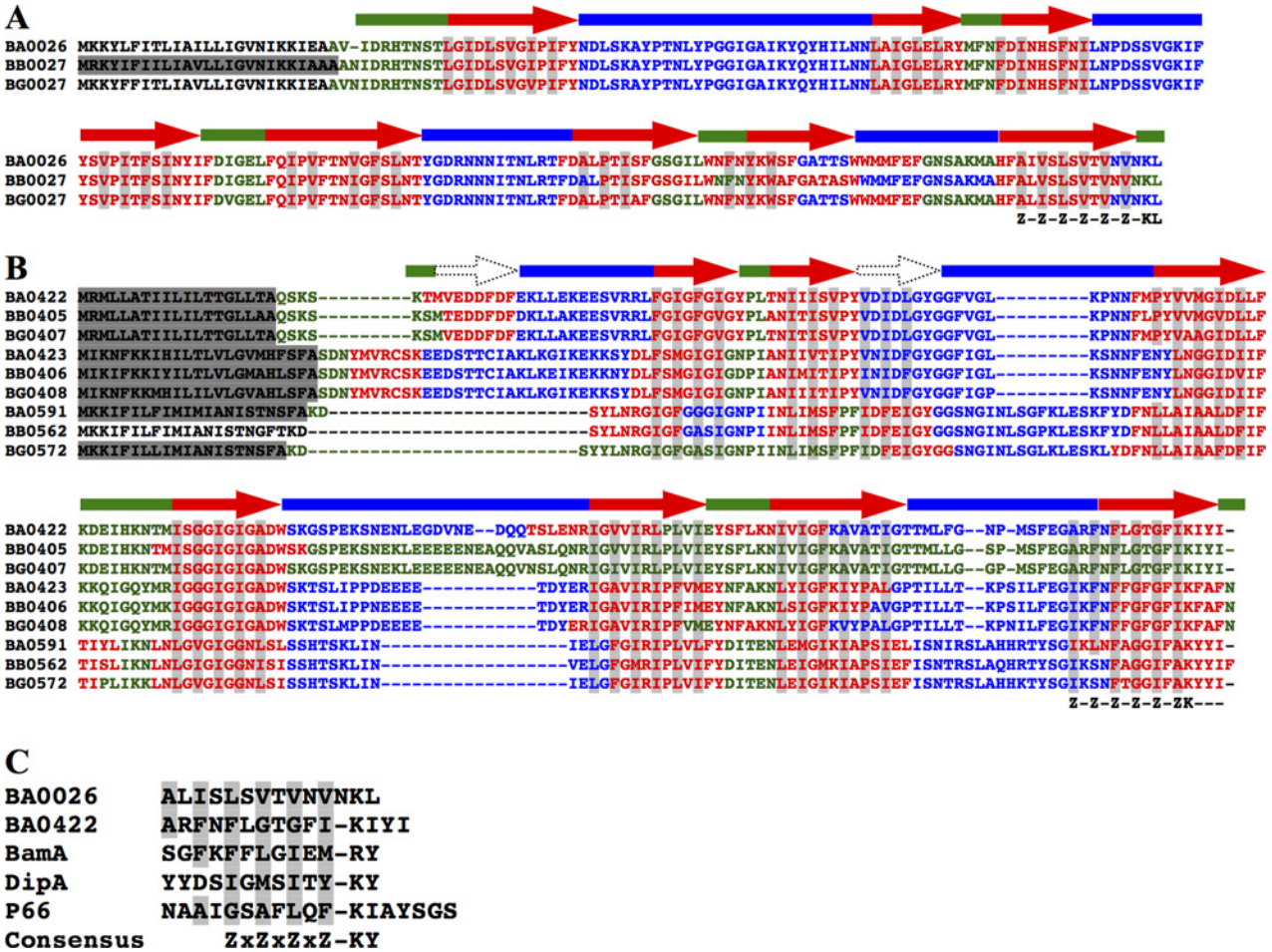


Figure 2 Multiple sequence alignment and consensus topology of the BAPKO_0026 group (A) and the BAPKO_0422 paralogous gene family (B)

Sequences from the three main Borrelia sub-species known to infect humans are included (BA, *B. afzelii*; *BB, B. burgdorferi*; and BG, *B. garinii*). Colours indicate predicted topology based on PRED-TMBB results (Figure 1). Non-polar membrane spanning residues are shaded light grey. Predicted N-terminal signal sequences are shaded dark grey (SignalP 4.1, using neural network 'SignalP-noTM'). Sequences were aligned using ClustalW. A consensus topology prediction is shown above each sequence alignment based on the PRED-TMBB topology and conserved hydrophobic residues. All sequences were analysed for potential stop-transfer sequences using ProtScale. The potential C-terminal signature sequence is indicated by Z-Z-Z-Z-Z-ZK. (**C**) An analysis of C-terminal signature sequences from experimentally confirmed β-barrels BamA (BB0795), DipA (BB0418) and P66 (BB_0603).

The four alternate non-polar residues are well conserved in the C-terminal region of the *Borrelia* OM-proteins (Figures 2A and 2B); however, some differences in the terminal residues are apparent. Only BB0562 terminates with a phenylalanine residue, almost all other sequences terminate with the small non-polar amino acids isoleucine or valine. The putative OM β-barrels identified in the present study contain an invariant lysine residue within four residues of the C-terminus. Following this lysine residue, the C-terminal sequence of the BAPKO_0422 group is rich tyrosine and small non-polar-residues. A comparison was made with the C-terminal regions of other experimentally confirmed *Borrelia* β-barrel proteins BamA, DipA and P66 (Figure 2C). The sequences reveal four alternate non-polar residues followed by a conserved positively charged residue (lysine or arginine).

Homology model of BAPKO_0422

To allow a more detailed analysis of residue packing within the barrel interior and positioning in the membrane, homology models were generated using Modeller [47] (Figure 3). This modelling benefited from the availability of several high-resolution structures of TM β-barrels from an evolutionary diverse range of bacteria including *E. coli*, *N. meningitidis*, *Y. pestis*, *Pseudomonas aeruginosa* and *Thermus thermophilus*. The four *B. afzelii* homologues shown in Figure 1 were used as target sequences. The models suggest that all four proteins have a vertical height 45–62 Å in the membrane with a width of ~25 Å (Figure 3). R_g values based on centre of mass were also calculated and ranged between 17.0 and 19 Å. The modelled β-barrels resemble an inverse micelle, with numerous polar residues and ion-pair interactions forming a tightly packed interior. This is surrounded by an

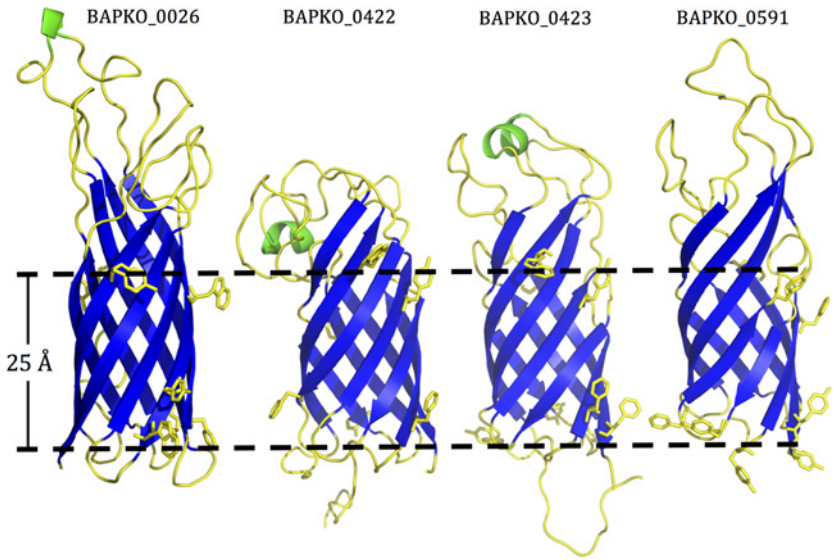


Figure 3 Homology models of BAPKO_0026, BAPKO_0422, BAPKO_0423 and BAPKO_0591 generated using Modeller
The predicted secondary structure shows an eight-stranded antiparallel β-barrel (blue). Loops and turns are shown in yellow, α-helices are shown in green. Aromatic-girdle residues are shown as sticks representation. The approximate limits of the aliphatic region are indicated by dashed lines. R_g values calculated in Chimera: BAPKO_0026 18.9 Å, BAPKO_0422 16.8 Å, BAPKO_0423 18.1 Å, BAPKO_0591 18.4 Å.

exterior band of aliphatic residues predicted to interact with the non-polar acyl chains of *Borrelia* glycolipids. Aromatic residues are frequently observed at the lipid-water interface in membrane spanning β-barrels of Gram-negative bacteria [26]. Two aromatic girdles are present in the *Borrelia* proteins, allowing the vertical positioning in the membrane to be estimated (dashed line, Figure 3). The distance between the two aromatic girdles (24–27 Å) is consistent with the hydrophobic distance (~26 Å) of β-barrels from other Gram-negative bacteria [58].

Expression, purification and characterization of recombinant BAPKO_0422

The predicted mature form of BAPKO_0422 (BAPKO_0422$^{20\text{-}201}$) was produced in the *E. coli* expression system with a cleavable N-terminal 6× His-tag and purified to homogeneity as described in 'Materials and Methods' (Figure 4A). The resultant protein lacked a functional N-terminal signal sequence and so was produced as inclusion bodies, facilitating separation from native *E. coli* membrane proteins. Purified BAPKO_0422$^{20\text{-}201}$ readily refolded using a variety of methods including dilution and on-column refolding protocols (see Materials and Methods). Numerous studies have demonstrated that eight-stranded β-barrel membrane proteins can spontaneously refold in lipid bilayers or detergent micelles following either heat- or chemically-induced unfolding [58].

The size exclusion chromatogram (Figure 4B) revealed a single peak. The elution volume corresponds to a molecular mass of 23 kDa indicating that BAPKO_0422 forms a monomer under the conditions tested. Protein folding was confirmed by CD, which revealed a classic β-strand-type spectrum with maxima at 196 nm and minima at 216 nm (Figure 5A). The secondary structure analysis predicts 40 % β-sheet, 20 % turn and 35 % unordered.

The relatively low cloud point of Triton X-114 (~23 °C) allows the separation of amphiphilic integral membrane proteins from water-soluble proteins at physiological temperatures. Above the cloud point, aggregation of detergent micelles results in a two-phase system consisting of an aqueous phase and a detergent-rich phase. This detergent phase can be readily isolated by low speed centrifugation. BAPKO_0422 was shown to be amphiphilic, partitioning to the detergent phase (Figure 5B).

Structural characterization of BAPKO_0422 by SAXS

SAXS data from a solution of a monodisperse protein can provide a measure of the R_g and overall molecular shape. SAXS data were recorded from both untagged BAPKO_0422$^{20\text{-}201}$ (Figure 6A) and His-tagged protein (6His-BAPKO_0422). Inspection of the Guinier region and residuals from linear fitting (Figure 6B) revealed linearity an indication of good quality data with minimal aggregation.

Radiation damage may result in sample aggregation or protein misfolding through the course of an experiment. Scattering data were compared from initial and final images with no significant changes observed in either scattering form or calculated R_g values. Samples were analysed by SDS/PAGE both before and after X-ray exposure and were unchanged. To test for the possibility of concentration dependent effects such as aggregation

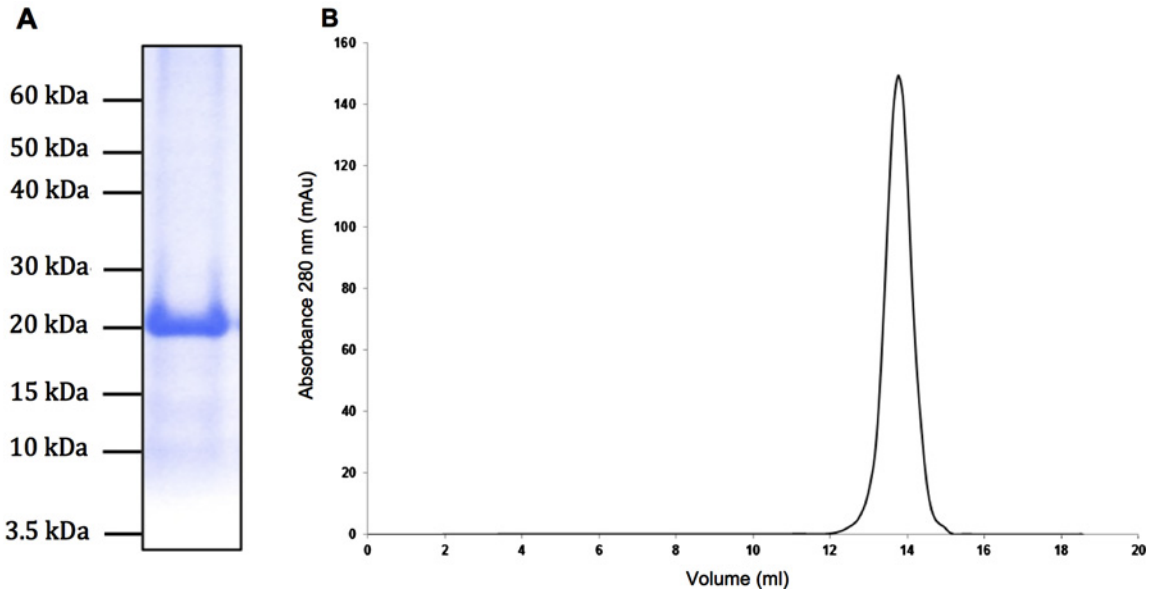


Figure 4 **Production and purification of recombinant BAPKO_0422$^{20\text{-}201}$**
(**A**) SDS/PAGE of purified BAPKO_0422$^{20\text{-}201}$ with a single band at ~23 kDa. (**B**) Size exclusion trace for BAPKO_0422$^{20\text{-}201}$. Purified protein was applied on to a Superdex 75 10/300 column at a flow rate of 1 ml/min resulting in a single sharp peak corresponding to an approximate molecular mass of 23 kDa.

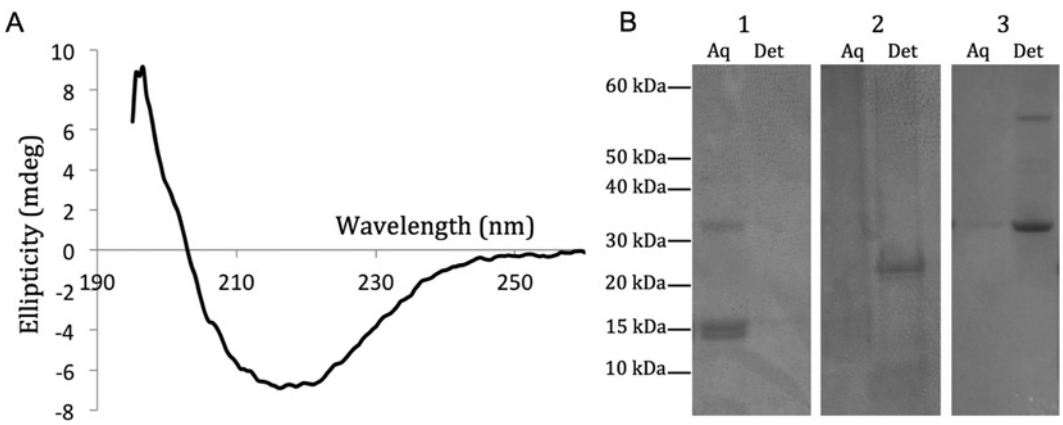


Figure 5 **CD (A) and phase partitioning (B) of BAPKO_0422$^{20\text{-}201}$**
CD data were acquired at 20 °C in duplicate at a protein concentration of 0.33 mg/ml in 0.1 % (w/v) DDAO, 0.3 M NaCl, 30 mM Tris/HCl, pH 8, between wavelengths of 195–260 nm. Maxima at 196 nm and minima at 216 nm indicate a structure rich in β-strand. Phase partition experiments were conducted using the non-ionic detergent Triton X-114, allowing separation of hydrophilic proteins (Aq) from amphiphillic detergent-soluble proteins (Det). Haemoglobin (1) was used as an aqueous phase negative control. BAPKO_0422$^{20\text{-}201}$(2) partitions to the detergent phase. AtVDAC1 (3) was used as a known detergent-phase positive control.

or inter-particle interference, R_g values were determined at three concentrations for both His-tagged and native protein (Table 2).

R_g values calculated by both Guinier analysis and Real Space analysis (Table 2) are comparable to the R_g calculated from the homology model (Figure 3) indicating that BAPKO_0422 is monomeric under the conditions tested. The pair-distance distribution function [P(r) function] gives information on the shape of a molecule by describing the paired-set of all distances between points in an object (Figure 6C). The single peak is indicative of a single-domain globular protein. The distribution of longest dimensions approaches zero with a concave slope and a D_{max} of 43 Å.

A Kratky plot [$q^2I(q)$ versus q] can be used to distinguish between compact folded structures and unfolded flexible systems. The data show a parabolic curve indicative of a compact globular structure (Figure 6D).

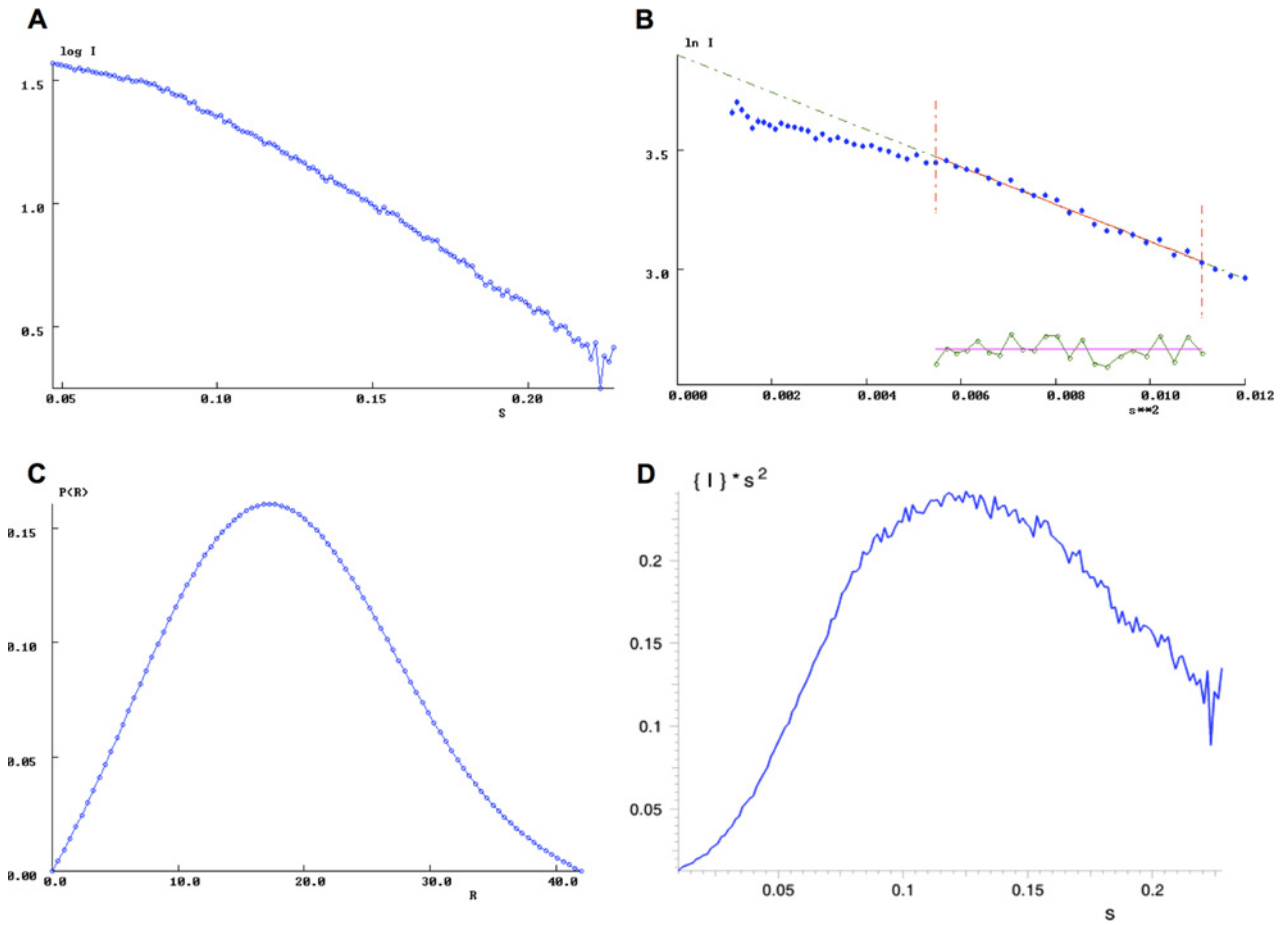


Figure 6 **SAXS data of BAPKO_0422$^{20\text{-}201}$**

Analysis of the scattering data generated from a sample of untagged BAPKO_0422$^{20\text{-}201}$ at 3.5 mg/ml in 0.3 M NaCl, 50 mM Tris-base, 0.1 % DDAO, pH 8.0. (**A**) The raw scattering data shown as a log of intensity over the scattering vector ($Å^{-1}$) and generated in Primus following background subtraction. (**B**) Guinier analysis of the background subtracted scattering data. A Guinier approximation was applied manually using Primus as ln[I(s)] versus s^2. The blue data points represent the scattering data, the red line shows the selected Guinier region and the green data illustrates the corresponding residuals. (**C**) The distance distribution function P(r) of BAPKO_0422$^{20\text{-}201}$ generated using GNOM with a D_{max} of 43 Å. (**D**) Kratky plot [$s^2 \times$ I(s) versus s] generated using Primus. The momentum transfer was defined as $s = 4\pi \sin(\theta)/\lambda$.

Table 2 **Calculated R_g values for BAPKO_0422$^{20\text{-}201}$**

A summary of R_g values for untagged and His-tagged BAPKO_0422$^{20\text{-}201}$ calculated using various methods. The R_g was calculated for both samples manually by Gunier approximation and then by Gnom in reciprocal and real space using the full curve.

	Untagged BAPKO_0422$^{20\text{-}201}$		
	2.5 mg/ml	**3.5 mg/ml**	**4.5 mg/ml**
Guinier R_g (Å)	13.5 ± 0.15	14.6 ± 0.11	12.5 ± 0.39
Reciprocal space R_g (Å)	15.7	14.3	14.4
Real space R_g (Å)	15.7	14.3	14.5
	His-tagged BAPKO_0422$^{20\text{-}201}$		
	3.0 mg/ml	**4.5 mg/ml**	**6.0 mg/ml**
Guinier R_g (Å)	14.2 ± 0.13	14.1 ± 0.12	14.0 ± 0.10
Reciprocal space R_g (Å)	16.8	16.8	16.6
Real space R_g (Å)	16.8	16.8	16.7

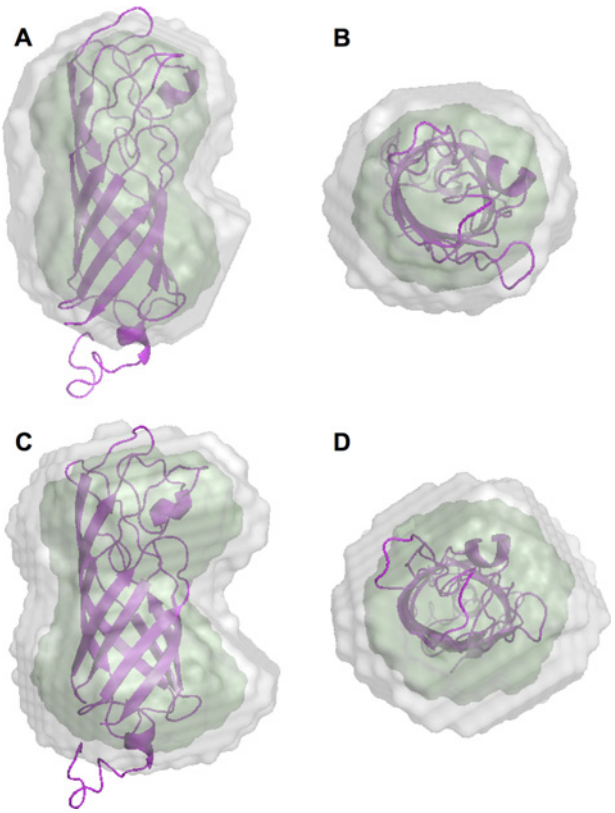


Figure 7 Low-resolution molecular envelopes of BAPKO_0422$^{20\text{-}201}$ (A and B) and His-tagged BAPKO_0422$^{20\text{-}201}$ (C and D) determined by SAXS

Refined and filtered molecular envelopes were generated from 20 independent DAMMIN models (Supplementary Figures S1 and S2). The green surface represents the filtered envelope generated by DAMFILT, the grey envelope is the computed probability map generated by DAMAVER. The homology model of BAPKO_0422$^{20\text{-}201}$ is shown as purple cartoon.

Ab initio molecular envelopes of native and His-tagged BAPKO_0422 calculated by simulated annealing reveal similar structures with some minor additional density in the latter (Figure 7). The SAXS envelope is consistent with the dimensions of a β-barrel and the homology model of BAPKO_0422 docked within the envelope shows close agreement (Figure 7).

BAPKO_0422 is an fH-binding protein

Numerous studies have revealed the complex differential binding of both human and animal fH by various strains of *Borrelia*, contributing significantly to pathogenicity and host-competence [59,60]. A broad screen of whole cell sonicate from *B. garinii* against human sera followed by *de novo* sequencing identified the hypothetical protein BG0407 (Genbank AAU07257) as a novel fH-binding protein [59]. To test the possibility that the close homologue BAPKO_0422 may also be an fH-binding protein we used a far western blot (ALBI assay). Briefly, immunoblots of BAPKO_0422 along with positive and negative controls were incubated with human fH. Bound fH was detected by a monoclonal Anti-human fH primary antibody, followed by a fluorescent secondary antibody (see Materials and Methods). The results demonstrated that recombinant BAPKO_0422 formed a specific interaction with human fH at a concentration 3-fold lower than that found in human blood (Figure 8).

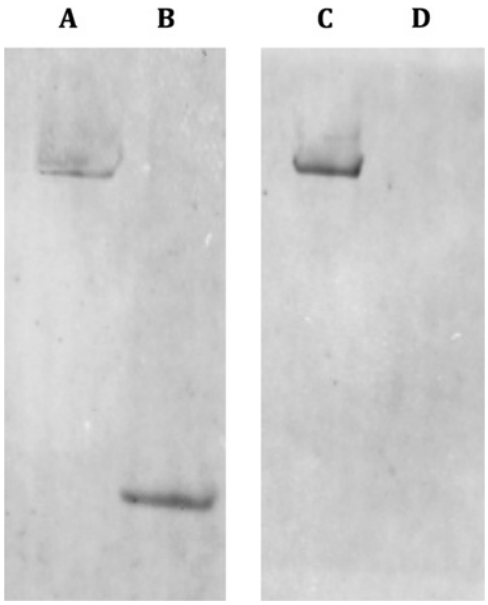


Figure 8 ALBI assay

Non-reducing 1D immunoblots of 0.5 μg of human fH as a positive control (lanes A and C) and 10 μg of recombinant BAPKO_0422$^{20\text{-}201}$ (lanes B and D). Blot 1 (A and B) was incubated with human fH (73 μg/ml), whereas blot 2 (C and D) was incubated with TBS buffer. Human fH was detected by a mouse monoclonal anti-human fH primary, followed by a fluorescent goat anti-mouse IgG secondary antibody, as described in 'Materials and Methods'.

DISCUSSION

Although the OM of *Borrelia* is distinct from Gram-negative bacteria, the central components of the Sec-dependent secretion pathway and the barrel assembly apparatus involved in localization and insertion of proteins into the OM (BamA, BamB, BamD, Skp), appear to be conserved [30,61]. Genome data are now available for numerous strains of *Borrelia*; however, it remains a significant challenge to identify all potential OM β-barrels due to the inherent difficulties in theoretical prediction. With the exception of the lipoproteins, very few surface exposed proteins have been identified in *Borrelia*. In the present study, we have identified a paralogous group of four *Borrelia* proteins which we predict are eight-stranded membrane-spanning β-barrels. As the most common cause of EM rash in Europe, we chose to study the four *B. afzelii* homologues in more detail. Bioinformatic analyses using a range of prediction methods indicate a topology similar

to the well-studied *E. coli* proteins OmpA, OmpX and OmpW. Experimental evidence from CD and a low-resolution SAXS structure of recombinant BAPKO_0422 support this hypothesis.

As recombinant protein was produced from the *E. coli* expression system and purified from inclusion bodies, confirmation of protein refolding is required. The CD data indicate recombinant BAPKO_0422$^{20\text{-}201}$ forms extensive secondary structure rich in β-strand. The results are directly comparable to a range of other β-barrels such as OmpW [62], OmpA and OprF [63]. In addition, Kratky plots generated from SAXS data indicate a compact, folded structure. The formation of tertiary structure of many β-barrels proteins can be monitored by the difference in apparent molecular mass between the folded and the unfolded states, as determined by SDS/PAGE. Numerous membrane-spanning β-barrels have been shown to maintain a folded state when solubilized in SDS at room temperature, but once denatured by boiling will remain in an unfolded state [64]. Therefore, gel-shift assays are conducted by comparing boiled and unboiled samples. Gel-shift assays were conducted using recombinant BAPKO_0422; however, the protein remained unmodified by heat (results not shown). This result is consistent with studies on the 24-stranded P66, which was also not modified by heat [28] and is perhaps indicative of a general feature of OMPs in *Borrelia*.

The *ab initio* molecular envelope of BAPKO_0422 determined by SAXS reveals a peanut shaped structure with dimensions 25 × 45 Å. This structure is consistent with a monomeric eight-stranded β-barrel and suggests a multimeric porin-type structure is unlikely. This is consistent with data on the orthologous protein from *B. burgdorferi* BB0405, which was shown to be monomeric while not making significant interactions with other major OM-proteins [65].

An analysis of N-terminal signal sequences of the BAPKO_0422 family of proteins revealed that the majority are predicted to have a functional signal sequence and so may enter the Sec-dependent secretory pathway for translocation across the inner membrane. In common with OM-proteins from other Gram-negative bacteria the *Borrelia* proteins identified in Table 1 are devoid of long hydrophobic stretches, precluding lateral transfer from the translocase to the inner membrane [66]. Some differences in N-terminal sequence between the three sub-species of *Borrelia* are seen. The N-terminal sequence of the *B. burgdorferi* protein BB0562 may not be recognized by signal peptidase I as it lacks a small non-polar residue normally present at the −1 position. Additionally BA0026 and BG0027 are not predicted to have functional signal sequences, whereas the close orthologue BB0027 does (Figure 2). These differences in predicted signal sequence may be due to subtle differences between the spirochetal signal peptides compared with those of the signal P training set. Alternatively, *Borrelia* is known to have three type-I signal peptidases compared with the single protein in *E. coli* [67]. This allows for the possibility of concerted divergence of signal peptides along with their corresponding peptidase resulting in a wider range of functional signal sequences in *Borrelia* compared with *E. coli*.

The homology models and topology predictions clearly show that the Borrelia OmpA-like domains listed in Table 1 are devoid of any C-terminal domains. The C-terminal residues are therefore predicted to form the terminal strand of the β-barrel and as such are expected to contain the highly conserved C-terminal signature sequence motif [68]. This 10-residue signature sequence consists of three alternating non-polar residues at positions −9, −7 and −5 from the C-terminus. These side chains make contact with the aliphatic region of the membrane. The −3 residue is usually a tyrosine residue that forms part of the aromatic girdle occupying the interface between polar and non-polar environments [57]. The C-terminal residue is usually a phenylalanine but may occasionally be substituted with other aromatic amino acids or more rarely other non-polar amino acids. Although not essential for proper processing, the aromatic nature of the C-terminal residue enhances processing by the BamA apparatus, facilitating barrel assembly in the OM [57,69]. Variations in the C-terminal signature sequence and their recognition by BamA have been shown to be species specific, in particular, many Proteobacterial OMPs have a positively charged residue in the penultimate position, whereas this is never observed in *E. coli* [70]. A C-terminal signature sequence is apparent among the BAPKO_0422 family and BA0026, both consisting of a series of six alternative non-polar residues, a single lysine residue and a small polar terminal motif (Figure 2). With the exception of BB0562, the most notable exception to the standard proteobacterial C-terminal sequence is the absence of a C-terminal phenylalanine residue. Instead the sequences terminate with either the small non-polar residues leucine, isoleucine or the polar asparagine. In addition, a number of tyrosine residues are also present in the final three residues. The differences observed in the present study may be suggestive of a *Borrelia*-specific C-terminal signature sequence. In order to confirm this, we analysed the C-terminal strands of the small number of experimentally confirmed β-barrels in *Borrelia*. DipA has a C-terminal signature sequence reminiscent of the BAKPO_0422 family, terminating with K-Y. In addition, the barrel assembly apparatus BamA, is itself a β-barrel and the last membrane-spanning strand terminates with R-Y. The terminal strand of the OM porin P66 also contains the conserved positive lysine residue and terminates with the polar sequence S-G-S. We therefore propose that these membrane-spanning β-barrels contain a *Borrelia*-specific C-terminal signature sequence and this is recognized efficiently by the *Borrelia* BamA apparatus. This sequence resembles Z-x-Z-x-Z-x-Z-[KR]-[ILNY], where Z represents a small non-polar residue, x can be any amino acid. A conserved positively charged residue is present, but is not necessarily the penultimate residue. The terminal residue appears to be variable and is rarely a phenylalanine residue. An analysis of *Borrelia* proteins identified as potential β-barrels by the TMBB-DataBase [71] reveals a range of C-terminal residues including isoleucine, valine, lysine, asparagine and tyrosine.

A literature search for homologues of BAPKO_0422 from other *Borrelia* subspecies revealed insights into possible functions and cellular localization. The *B. burgdorferi* homologue BB0405 was shown to be surface exposed, present in OM vesicles [65] and expressed in conditions representative of both tick and host environments [72]. Another homologue, BB0407 (Genbank AAU07257) was identified from a screen of proteins with

fH-binding activity suggesting a possible role in virulence [59]. This fH-binding activity led us to investigate the fH binding of BAPKO_0422, which we confirmed using an ALBI assay. These results and the high sequence similarity suggest that the orthologous group BAPKO_0422, BG0407 and BB0405 are expressed in the mammalian host, exposed at the cell surface and bind to human fH.

In summary, the data presented suggest that BAPKO_0422 forms a membrane-spanning β-barrel. The topology prediction matches the OmpA-membrane spanning domain defined by Pfam family PF01389, consisting of eight membrane-spanning β-strands linked by short periplasmic turns and longer extracellular loops. For the first time, this extends the species distribution of the PF01389 domain to include members of the *Spirochete* phylum. The orthologous group consisting of BAPKO_0422, BB0405 and BG0407 are proposed to play a role in virulence by binding to host fH, therefore abrogating the host complement response and reducing antigenicity of surface exposed loops. Further work is required to establish the host-specific fH-binding activity of the remaining homologues and to determine any other physiological functions of these proteins in *Borrelia.* Surface exposed epitopes may enhance recombinant immunoblots currently used in diagnostic tests for Lyme disease.

AUTHOR CONTRIBUTION

Adam Dyer, Gemma Brown, Lenka Stejskal and Peter Laity designed and conducted experiments and performed data analysis. Richard Bingham designed the study and supervised the research. All authors participated in writing the paper.

ACKNOWLEDGEMENTS

We are grateful to the following: Ibad Kureshi and members of the HPC-RC team at the University of Huddersfield for computational support, Andrew Leech for CD data collection, Gabriele Margos, Stephanie Vollmer, Ruth Mitchell and Freddie Seelig for tick collection and DNA extraction and to George Psakis and Alexandre Boulbrima for provision of AtVDAC1.

FUNDING

This work was supported by the Biochemical Society (to L.S.); and the University of Huddersfield, Department of Biological Sciences.

REFERENCES

1 Steere, A.C., Coburn, J. and Glickstein, L. (2004) The emergence of Lyme disease. J. Clin. Invest. **113**, 1093–1101 CrossRef PubMed

2 Margos, G., Vollmer, S.A., Ogden, N.H. and Fish, D. (2011) Population genetics, taxonomy, phylogeny and evolution of Borrelia burgdorferi sensu lato. Infect. Genet. Evol. **11**, 1545–1563 CrossRef PubMed

3 Ornstein, K., Berglund, J., Nilsson, I., Norrby, R. and Bergstrom, S. (2001) Characterization of Lyme borreliosis isolates from patients with erythema migrans and neuroborreliosis in southern Sweden. J. Clin. Microbiol. **39**, 1294–1298 CrossRef PubMed

4 Oksi, J., Marjamaki, M., Nikoskelainen, J. and Viljanen, M.K. (1999) *Borrelia burgdorferi* detected by culture and PCR in clinical relapse of disseminated Lyme borreliosis. Ann. Med. **31**, 225–232 CrossRef PubMed

5 Schmidli, J., Hunziker, T., Moesli, P. and Schaad, U.B. (1988) Cultivation of *Borrelia burgdorferi* from joint fluid three months after treatment of facial palsy due to Lyme borreliosis. J. Infect. Dis. **158**, 905–906 CrossRef PubMed

6 Preac-Mursic, V., Weber, K., Pfister, H.W., Wilske, B., Gross, B., Baumann, A. and Prokop, J. (1989) Survival of Borrelia burgdorferi in antibiotically treated patients with Lyme borreliosis. Infection **17**, 355–359 CrossRef PubMed

7 Feng, J., Wang, T., Shi, W., Zhang, S., Sullivan, D., Auwaerter, P.G. and Zhang, Y. (2014) Identification of novel activity against Borrelia burgdorferi persisters using an FDA approved drug library. Emerg. Microbes Infect. **3**, e49 CrossRef PubMed

8 Coutte, L., Botkin, D.J., Gao, L. and Norris, S.J. (2009) Detailed analysis of sequence changes occurring during vlsE antigenic variation in the mouse model of *Borrelia burgdorferi* infection. PLoS Pathog. **5**, e1000293 CrossRef PubMed

9 Rogovskyy, A.S. and Bankhead, T. (2013) Variable VlsE is critical for host reinfection by the Lyme disease spirochete. PLoS One **8**, e61226 CrossRef PubMed

10 Tilly, K., Bestor, A. and Rosa, P.A. (2013) Lipoprotein succession in *Borrelia burgdorferi*: similar but distinct roles for OspC and VlsE at different stages of mammalian infection. Mol. Microbiol. **89**, 216–227 CrossRef PubMed

11 Zhang, J.R., Hardham, J.M., Barbour, A.G. and Norris, S.J. (1997) Antigenic variation in Lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. Cell **89**, 275–285 CrossRef PubMed

12 Crother, T.R., Champion, C.I., Whitelegge, J.P., Aguilera, R., Wu, X.Y., Blanco, D.R., Miller, J.N. and Lovett, M.A. (2004) Temporal analysis of the antigenic composition of *Borrelia burgdorferi* during infection in rabbit skin. Infect. Immun. **72**, 5063–5072

13 Dietrich, T., Geissdorfer, W., Schlotzer-Schrehardt, U., Holbach, L., Schoerner, C. and Seitz, B. (2008) Borrelia-associated crystalline keratopathy with intracorneal detection of *Borrelia garinii* by electron microscopy and polymerase chain reaction. Cornea **27**, 498–500 CrossRef PubMed

14 Livengood, J.A. and Gilmore, R.D. (2006) Invasion of human neuronal and glial cells by an infectious strain of *Borrelia burgdorferi*. Microbes Infect. **8**, 2832–2840 CrossRef PubMed

15 Pulzova, L. and Bhide, M. (2014) Outer surface proteins of Borrelia: peerless immune evasion tools. Curr. Protein Pept. Sci. **15**, 75–88 CrossRef PubMed

16 Kurtenbach, K., de Michelis, S., Etti, S., Schafer, S.M., Sewell, H.S., Brade, V. and Kraiczy, P. (2002) Host association of *Borrelia burgdorferi* sensu lato–the key role of host complement. Trends Microbiol. **10**, 74–79 CrossRef PubMed

17 Behera, A.K., Durand, E., Cugini, C., Antonara, S., Bourassa, L., Hildebrand, E., Hu, L.T. and Coburn, J. (2008) *Borrelia burgdorferi* BBB07 interaction with integrin alpha3beta1 stimulates production of pro-inflammatory mediators in primary human chondrocytes. Cell Microbiol. **10**, 320–331 PubMed

18 Ostberg, Y., Pinne, M., Benz, R., Rosa, P. and Bergstrom, S. (2002) Elimination of channel-forming activity by insertional inactivation of the p13 gene in *Borrelia burgdorferi*. J. Bacteriol. **184**, 6811–6819 CrossRef PubMed

19 Barcena-Uribarri, I., Thein, M., Barbot, M., Sans-Serramitjana, E., Bonde, M., Mentele, R., Lottspeich, F., Bergström, S. and Benz, R. (2014) Study of the protein complex, pore diameter, and pore-forming activity of the *Borrelia burgdorferi* P13 porin. J. Biol. Chem. **289**, 18614–18624 CrossRef PubMed

20 Bunikis, I., Denker, K., Ostberg, Y., Andersen, C., Benz, R. and Bergstrom, S. (2008) An RND-type efflux system in Borrelia burgdorferi is involved in virulence and resistance to antimicrobial compounds. PLoS Pathog **4**, e1000009 CrossRef PubMed

21 Greene, N.P., Hinchliffe, P., Crow, A., Ababou, A., Hughes, C. and Koronakis, V. (2013) Structure of an atypical periplasmic adaptor from a multidrug efflux pump of the spirochete *Borrelia burgdorferi*. FEBS Lett. **587**, 2984–2988 CrossRef PubMed

22 Setubal, J.C., Reis, M., Matsunaga, J. and Haake, D.A. (2006) Lipoprotein computational prediction in spirochaetal genomes. Microbiology **152**, 113–121 CrossRef PubMed

23 Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., White, O., Ketchum, K.A., Dodson, R. and Hickey, E.K. (1997) Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature **390**, 580–586 CrossRef PubMed

24 Casjens, S.R., Mongodin, E.F., Qiu, W.G., Dunn, J.J., Luft, B.J., Fraser-Liggett, C.M. and Schutzer, S.E. (2011) Whole-genome sequences of two *Borrelia afzelii* and two *Borrelia garinii* Lyme disease agent isolates. J. Bacteriol. **193**, 6995–6996 CrossRef PubMed

25 Cabello, F.C., Godfrey, H.P. and Newman, S.A. (2007) Hidden in plain sight: *Borrelia burgdorferi* and the extracellular matrix. Trends Microbiol. **15**, 350–354 CrossRef PubMed

26 Schulz, G.E. (2000) Beta-barrel membrane proteins. Curr. Opin. Struct. Biol. **10**, 443–447 CrossRef PubMed

27 Thein, M., Bonde, M., Bunikis, I., Denker, K., Sickmann, A., Bergstrom, S. and Benz, R. (2012) DipA, a pore-forming protein in the outer membrane of Lyme disease spirochetes exhibits specificity for the permeation of dicarboxylates. PLoS One **7**, e36523 CrossRef PubMed

28 Kenedy, M.R., Luthra, A., Anand, A., Dunn, J.P., Radolf, J.D. and Akins, D.R. (2013) Structural modeling and physicochemical characterization provide evidence that P66 forms a beta-barrel in the *Borrelia burgdorferi* outer membrane. J. Bacteriol. **196**, 859–872 CrossRef PubMed

29 Skare, J.T., Mirzabekov, T.A., Shang, E.S., Blanco, D.R., Erdjument-Bromage, H., Bunikis, J., Bergström, S., Tempst, P., Kagan, B.L., Miller, J.N. and Lovett, M.A. (1997) The Oms66 (p66) protein is a *Borrelia burgdorferi* porin. Infect. Immun. **65**, 3654–3661 PubMed

30 Lenhart, T.R. and Akins, D.R. (2010) *Borrelia burgdorferi* locus BB0795 encodes a BamA orthologue required for growth and efficient localization of outer membrane proteins. Mol. Microbiol. **75**, 692–709 CrossRef PubMed

31 Yamashita, S., Lukacik, P., Barnard, T.J., Noinaj, N., Felek, S., Tsang, T.M., Krukonis, E.S., Hinnebusch, B.J. and Buchanan, S.K. (2011) Structural insights into Ail-mediated adhesion in *Yersinia pestis*. Structure **19**, 1672–1682 CrossRef PubMed

32 Hong, H.D., Patel, D.R., Tamm, L.K. and van den Berg, B. (2006) The outer membrane protein OmpW forms an eight-stranded beta-barrel with a hydrophobic channel. J. Biol. Chem. **281**, 7568–7577 CrossRef PubMed

33 Vogt, J. and Schulz, G.E. (1999) The structure of the outer membrane protein OmpX from *Escherichia coli* reveals possible mechanisms of virulence. Structure **7**, 1301–1309 CrossRef PubMed

34 Pautsch, A. and Schulz, G.E. (1998) Structure of the outer membrane protein A transmembrane domain. Nat. Struct. Biol. **5**, 1013–1017 CrossRef PubMed

35 Koebnik, R., Locher, K.P. and van Gelder, P. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol. Microbiol. **37**, 239–253 CrossRef PubMed

36 Mittal, R., Krishnan, S., Gonzalez-Gomez, I. and Prasadarao, N.V. (2011) Deciphering the roles of outer membrane protein A extracellular loops in the pathogenesis of *Escherichia coli* K1 meningitis. J. Biol. Chem. **286**, 2183–2193 CrossRef PubMed

37 Prasadarao, N.V. (2002) Identification of *Escherichia coli* outer membrane protein A receptor on human brain microvascular endothelial cells. Infect. Immun. **70**, 4556–4563

38 Nair, M.K., Venkitanarayanan, K., Silbart, L.K. and Kim, K.S. (2009) Outer membrane protein A (OmpA) of Cronobacter sakazakii binds fibronectin and contributes to invasion of human brain microvascular endothelial cells. Foodborne Pathog. Dis. **6**, 495–501 CrossRef PubMed

39 Heffernan, E.J., Harwood, J., Fierer, J. and Guiney, D. (1992) The *Salmonella typhimurium* virulence plasmid complement resistance gene rck is homologous to a family of virulence-related outer membrane protein genes, including pagC and ail. J. Bacteriol. **174**, 84–91 PubMed

40 Crago, A.M. and Koronakis, V. (1999) Binding of extracellular matrix laminin to *Escherichia coli* expressing the Salmonella outer membrane proteins Rck and PagC. FEMS Microbiol. Lett. **176**, 495–501 CrossRef PubMed

41 Eddy, S.R. (1998) Profile hidden Markov models. Bioinformatics **14**, 755–763 CrossRef PubMed

42 Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J. et al. (2014) Pfam: the protein families database. Nucleic. Acids Res. **42**, D222–230 CrossRef PubMed

43 Sperisen, P. and Pagni, M. (2005) JACOP: a simple and robust method for the automated classification of protein sequences with modular architecture. BMC Bioinformatics **6**, 216 CrossRef PubMed

44 Petersen, T.N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods **8**, 785–786 CrossRef PubMed

45 Jaroszewski, L., Li, Z., Cai, X.H., Weber, C. and Godzik, A. (2011) FFAS server: novel features and applications. Nucleic. Acids. Res. **39**, W38–44 CrossRef PubMed

46 Bagos, P.G., Liakopoulos, T.D., Spyropoulos, I.C. and Hamodrakas, S.J. (2004) A Hidden Markov Model method, capable of predicting and discriminating beta-barrel outer membrane proteins. BMC Bioinformatics **5**, 29 CrossRef PubMed

47 Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U. and Sali, A. (2007) Comparative protein structure modeling using MODELLER. Curr. Protoc. Protein Sci. **Chapter 2**, Unit 2 9

48 Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. **25**, 1605–1612 CrossRef PubMed

49 Kopp, A., Hebecker, M., Svobodova, E. and Jozsi, M. (2012) Factor h: a complement regulator in health and disease, and a mediator of cellular interactions. Biomolecules **2**, 46–75 CrossRef PubMed

50 Bordier, C. (1981) Phase separation of integral membrane proteins in Triton X-114 solution. J. Biol. Chem. **256**, 1604–1607 PubMed

51 Mertins, B., Psakis, G., Grosse, W., Back, K.C., Salisowski, A., Reiss, P., Koert, U. and Essen, L.O. (2012) Flexibility of the N-terminal mVDAC1 segment controls the channel's gating behavior. PLoS One **7**, e47938 CrossRef PubMed

52 Volkov, V.V. and Svergun, D.I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. J. Appl. Cryst. **36**, 860–864 CrossRef

53 Svergun, D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. **76**, 2879–2886 CrossRef PubMed

54 Di, L., Pagan, P.E., Packer, D., Martin, C.L., Akther, S., Ramrattan, G., Mongodin, E.F, Fraser, C.M., Schutzer, S.E., Luft, B.J. et al. (2014) BorreliaBase: a phylogeny-centered browser of Borrelia genomes. BMC Bioinformatics **15**, 233 CrossRef PubMed

55 Cullen, P.A., Haake, D.A. and Adler, B. (2004) Outer membrane proteins of pathogenic spirochetes. FEMS Microbiol Rev **28**, 291–318 CrossRef PubMed

56 Pugsley, A.P. (1993) The complete general secretory pathway in gram-negative bacteria. Microbiol. Rev. **57**, 50–108 PubMed

57 Struyve, M., Moons, M. and Tommassen, J. (1991) Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J. Mol. Biol. **218**, 141–148 CrossRef PubMed

58 Tamm, L.K., Hong, H. and Liang, B. (2004) Folding and assembly of beta-barrel membrane proteins. Biochim. Biophys. Acta **1666**, 250–263 CrossRef PubMed

59 Bhide, M.R., Escudero, R., Camafeita, E., Gil, H., Jado, I. and Anda, P. (2009) Complement factor H binding by different Lyme disease and relapsing fever Borrelia in animals and human. BMC Res. Notes **2**, 134 CrossRef PubMed

60 Kraiczy, P. and 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 CrossRef PubMed

61 Guina, T., Helfet-Hilliker, D., Ramamurthy, V. and Oliver, D. (1998) Sequence and phylogenetic analysis of the *Borrelia burgdorferi* secA gene. Biochim. Biophys. Acta **1371**, 24–30 CrossRef PubMed

62 Albrecht, R., Zeth, K., Soding, J., Lupas, A. and Linke, D. (2006) Expression, crystallization and preliminary X-ray crystallographic studies of the outer membrane protein OmpW from *Escherichia coli*. Acta. Crystallogr. Sect. F Struct. Biol. Cryst. Commun. **62**, 415–418 CrossRef PubMed

63 Sugawara, E., Steiert, M., Rouhani, S. and Nikaido, H. (1996) Secondary structure of the outer membrane proteins OmpA of *Escherichia coli* and OprF of *Pseudomonas aeruginosa*. J. Bacteriol. **178**, 6067–6069 PubMed

64 Fairman, J.W., Noinaj, N. and Buchanan, S.K. (2011) The structural biology of beta-barrel membrane proteins: a summary of recent reports. Curr. Opin. Struct. Biol. **21**, 523–531 CrossRef PubMed

65 Yang, X., Promnares, K., Qin, J., He, M., Shroder, D.Y., Kariu, T., Wang, Y. and Pal, U. (2011) Characterization of multiprotein complexes of the *Borrelia burgdorferi* outer membrane vesicles. J. Proteome Res. **10**, 4556–4566 CrossRef PubMed

66 Xie, K., Hessa, T., Seppala, S., Rapp, M., von Heijne, G. and Dalbey, R.E. (2007) Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. Biochemistry **46**, 15153–15161 CrossRef PubMed

67 Paetzel, M., Dalbey, R.E. and Strynadka, N.C. (2000) The structure and mechanism of bacterial type I signal peptidases. A novel antibiotic target. Pharmacol. Ther. **87**, 27–49

68 Tommassen, J. (2010) Assembly of outer-membrane proteins in bacteria and mitochondria. Microbiology **156**, 2587–2596 CrossRef PubMed

69 Gessmann, D., Chung, Y.H., Danoff, E.J., Plummer, A.M., Sandlin, C.W., Zaccai, N.R. and Fleming, K.G. (2014) Outer membrane beta-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. Proc. Natl. Acad. Sci. U.S.A. **111**, 5878–5883 CrossRef PubMed

70 Robert, V., Volokhina, E.B., Senf, F., Bos, M.P., Van Gelder, P. and Tommassen, J. (2006) Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. PLoS Biol **4**, e377 CrossRef PubMed

71 Freeman, Jr, T.C. and Wimley, W.C. (2012) TMBB-DB: a transmembrane beta-barrel proteome database. Bioinformatics **28**, 2425–2430 CrossRef PubMed

72 Brooks, C.S., Vuppala, S.R., Jett, A.M. and Akins, D.R. (2006) Identification of *Borrelia burgdorferi* outer surface proteins. Infect. Immun. **74**, 296–304 CrossRef PubMed

Received 27 April 2015/23 June 2015; accepted 6 July 2015

Accepted Manuscript online 9 July 2015, doi 10.1042/BSR20150095