Summary

Despite major advances in the field of meningococcal pathogenesis research in recent decades, *Neisseria meningitidis* still remains a leading cause of meningitis and septicaemia world-wide. The high mortality and morbidity rates of meningococcal infection drive further research towards identifying novel and more effective therapeutic measures.

This study focuses on a small secreted protein found to have haem binding activity and an ability to interact with red blood cells. Gly1ORF1 protein is highly conserved and ubiquitous among the invasive strains of *N. meningitidis*. The crystal structure of the Gly1ORF1 K12A variant, solved in this study, showed that the protein possess a novel fold and is present in solution as a dimer. This study provides the first quantitative determination of Gly1ORF1 – haem interaction and has enabled the estimation of the dissociation constant for Gly1ORF1 variants.

Gly1ORF1's role in iron acquisition was suggested by a previous study. Therefore, the haemoglobin receptor HmbR was expressed in *E. coli* and was tested for the interaction with Gly1ORF1. The preliminary data suggests that there is interaction between those proteins. Additionally, this study shows that Gly1ORF1 is immunogenic in man and animals and that antibodies against this protein are able to direct complement-mediated killing of *N. meningitidis.*

A newly identified group of Gly1ORF1 homologs was also studied in this work. Two of the potential Gly1ORF1 homologs and one neisserial paralog were expressed in *E. coli* and subjected to functional analysis. The data suggest that the Gly1ORF1 homolog protein family is a group of small secreted proteins with shared secondary structure characteristics, but may have diverse functions. They might constitute a novel family of the bacterial virulence factors that could be novel therapeutic targets.

Acknowledgments

First and foremost I would like to thank my supervisor Professor Jon Sayers for all the guidance, help, and understanding throughout my PhD studies at the University of Sheffield. This thesis would not be completed without his encouragement and support. I appreciate his contribution of ideas and time – especially in the final months of my work.

I would like to take this opportunity to thank senior members of my group: Dr Sarbendra Pradhananga and Dr Guta Vitovski for their guidance and help throughout my project. I also thank Janine Phipps for her technical support and fulfilling all kind of request that have made my experimental work much easier.

Next, I thank all members of Sayer's group with whom I shared my everyday working life. I am really grateful to Dr Sudharsan Sathyamurthy, Dr Jing Zhang and Dr Ionam Wong for the help and friendship in the first years of my PhD. I also thank new members of my group: Sarah Oates, Hannah McMellon, Sam Harding, Roxanne Lau and Jess Tarrant for enjoyable atmosphere in the lab and constant help. I am especially grateful to Sarb, Sarah and Hannah for their time and effort they have spent proofreading this thesis. I also thank all members of Dr Mark Thomas and Dr Jon Shaw's groups for providing me with very useful advice and constructive feedback on my research.

I would also like to acknowledge members of Professor Artymiuk's group: Jason Wilson and Dr Rick Salmon for performing protein crystallisation trials. I am also really grateful to Dr Myron Christodoulides and members of his group: Miao Chiu Hung and Magdalena Bielecka for their advice and help with SBA experiments.

Finally, I would like to express my special thanks to my husband Rafal and my daughter Helena for their constant support and encouragement all the way through. I would not be able to get to this point without their unconditional love, patience and understanding. I also thank my parents and my closest family members for believing in me and for looking after me and my daughter in the final stages of writing up.

List of Contents

List of Figures

List of Tables

List of Abbreviations

°C	Degree Celsius
µM	Micromol
ABTS	2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
ACP	Adhesin complex protein
APS	Ammonium persulphate
ATP	Adenosine 5'-triphosphate
AU	Absorbance units
BBB	Blood Brain Barrier
BER	Base excision repair
Bicine	N,N-bis[2-hydroxyethyl]glycine
bp	Base pair
BSA	Bovine serum albumin
CAECAM	Carcinoembryonic antigen-related cell-adhesion molecule
CBA	Colombia blood agar
CFU	Colony forming units
CNBr	Cyanogen bromide
CNS	Central nervous system
CSF	Cerebrospinal fluid
DEFRA	Department for Environment, Food & Rural Affairs
DIC	Disseminated intravascular coagulopathy
DMSO	Dimethyl sulphoxide
DNA	Deoxy ribonucleic acid
dNTP	Deoxynucleotide triphosphate
ds	Double-stranded
DTBP	Dimethyl 3,3'-dithiobispropionimidate*2HCl

DTSSP	3,3'-dithiobis[sulfosuccinimidylpropionate]
DTT	Dithiothreitol
EBP	Erythrocyte band protein
EDTA	Ethylene diamine tetra acetic acid
ELISA	Enzyme-linked immunoabsorbent assay
FACS	Fluorescence-activated cell sorting
Fbp	Ferric binding protein
FCS	Foetal calf serum
fH	Factor H
fHBP	Factor H binding protein
FITC	Fluorescein isothiocyanate
Gly1	Gonolysin 1
GuHCl	Guanidine hydrochloride
Hb	Haemoglobin
HC HI	Heat-inactivated human complement
HC	Human complement
HCl	Hydrochloric acid
HRP	Horseradish Peroxidase
HSPG	Heparin sulphate peptidoglycan
IgA	Immunoglobulin A
IL	Interleukin
IMAC	Metal ion affinity chromatography
IPTG	Isopropyl-D-1-thiogalactopyranoside
Kb	Kilo base
K_d	Dissociation constant
kDa	Kilo Dalton
KO	Knock-out

LAMP	Lysosomal associated membrane protein
LB	Luria-Bertani
Lbp	Lactoferrin binding protein
Lf	Lactoferrin
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
MAC	Membrane attack complex
mg	Milligramme
ml	Millilitre
mM	Millimol
NadA	Neisseria adhesin A
NaOAc	Sodium Acetate
ng	Nanogramme
NG	*Neisseria gonorrhoeae*
Ni-affinity	Nickel affinity
NL	*Neisseria lactamica*
nM	Nanomol
NM	*Neisseria meningitidis*
NTPs	nucleotide triphosphates
OD	Optical density
OMV	outer membrane vesicle
ORF	Open reading frame
PAGE	Polycrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEI	Polyethyleneimine

PMSF	Phenylmethanesulfonyl fluoride
PNK	Polynucleotide kinase
RBCs	Red blood cells
RNA	Ribonucleic acid
RNase	Ribonuclease
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SEM	The standard error of the mean
SEU	Serum Epidemiological Unit
ss	Single-stranded
TBDT	TonB dependent transporter
Tbp	Transferrin binding protein
TEMED	N,N, N',N'-tetramethylethylenediamine
Tf	Transferrin
TNF	Tumour necrosis factor
Tris	Tris[hydroxymethyl]aminomethane
UK	United Kingdom
UV	Ultra-violet
WT	Wild type
YT	Yeast tryptone
µl	Microlitre

CHAPTER 1   INTRODUCTION

1.1 Epidemiology of *Neisseriae*

Neisseria meningitidis is an obligate human pathogen of the family *Neisseriaceae*. The Gram-negative diplococci colonise the human upper respiratory tract asymptomatically, with preference for the epithelium of nasopharynx. The global carriage rate is estimated to be 10-35%. However, in some populations (i.e. university students and military recruits) it can reach up to 100% (Hill *et al.*, 2010). Despite the high disease incidence rate in infants, carriage is more prevalent in young adults (Rosenstein *et al.*, 2001). Higher carriage rate is associated with decreased resistance of the mucosal barrier in the nasopharynx which can be observed during the dry season or following flu infection. Also, smoking, exposure to dust and overcrowding were identified as contributing factors to meningococcal colonisation (Stanwell-Smith *et al.*, 1994; Wilder-Smith *et al.*, 2002; Yazdankhah and Caugant, 2004).

It is not completely understood what triggers the transformation of the carriage strains into the invasive state, resulting in systemic invasion of the human blood stream. The main clinically important feature of *Neisseria meningitidis* transformation to the invasive state is the ability to rapidly multiply in the blood stream. This causes the acute form of septic shock known as *purpura fulminans,* which is associated with leakage of the vascular epithelium, disseminated intravascular coagulopathy (DIC) and organ failure. Once in the blood stream, meningocci can cross the blood brain barrier and multiply uncontrollably in the cerebrospinal fluid. This is manifested as a central nervous system disease – meningitis (Coureuil *et al.*, 2013). The similarity of the symptoms to the viral infections and rapid progression of the disease provide a challenge in the treatment of meningococcemia. Despite available antibiotic therapies, the mortality rate across Europe is 8% and higher in developing countries. The most serious outcomes are associated with meningococcal fulminant septicaemia (55%) and meningitis with septicaemia (25%), with lowest mortality rate for meningitis without associated septicaemia (generally 5%) (Trotter *et al.*, 2007b). Epidemic outbreaks often caused by hypervirulent strains

can reach even higher mortality rates. For example, in Norway between 1985-2002, the case fatality ratios (CFRs) of the meningococcal disease caused by serogroup C strains C:15:P1.7,16/ST-32 and C:2a/ST-11 was 21.1% and 18.2% respectively (Smith *et al.*, 2006). A Recently described outbreak in Mexico, caused by ST-11 lineage, was associated with a 36.8% mortality rate (Chacon-Cruz *et al.*, 2014). The survivors are often left with devastating complications including limb amputations, hearing loss and neurological disorders (Wong *et al.*, 2007).

The incidence of sporadic cases of meningococcemia varies between geographic regions, with <2 cases/100000 population *per annum* in European countries, United States and China and regions with more than 10 cases per 100000 population *per annum* in African countries (Jafri *et al.*, 2013). In the sub-Saharan African meningitis belt, seasonal outbreaks of the disease are observed yearly during the dry season. In addition, cyclical pandemics occur every 8-10 years, with the incidence rate over >1/1000 population, despite the introduction of the meningococcal vaccines (Stephens, 2009). In temperate climates, a higher incidence rate of meningococcal disease is observed during winters, with the highest peaks two weeks after influenza outbreaks (Cartwright *et al.*, 1991).

Neisseria meningitidis strains are classified into 13 serogroups based on the capsule polysaccharide (A, B, C, D, 29E, H, I, K, Y, W-135, X, Z) (Cartwright *et al.*, 1991). Meningococcal diseases are caused mainly by serotypes A, B, C, Y, W-135, with the geographical variation of prevalence for each serotypes. The majority of the meningococcal infection cases in Europe and America arise due to serotypes B and C, whereas in the meningococcal belt in Africa the majority of the infections are associated with serotypes A and C. Serotype W-135 is responsible for diseases worldwide with no regional differences in prevalence, whereas serotype Y was mainly reported in Canada (Stephens, 2009). In 1991, New Zealand was affected by a meningitis epidemic dominated by serogroup B strain B:4:P1.7b,4 (NZ 98/254) (Wong *et al.*, 2007). The introduction of the MeNZB vaccine, based on outer membrane vesicles of this strain, proved to be effective in controlling the New Zealand epidemic (Arnold *et al.*, 2011).

1.2 Pathogenic mechanisms to overcome human defences

The pathogenesis of meningococci can be divided into five stages (Fig. 1.1):

1. Colonisation of the nasopharyngeal mucosa
2. Invasion of the epithelium
3. Invasion of the blood stream
4. Crossing of the blood-brain barrier
5. Survival in the cerebrospinal fluid (CSF)
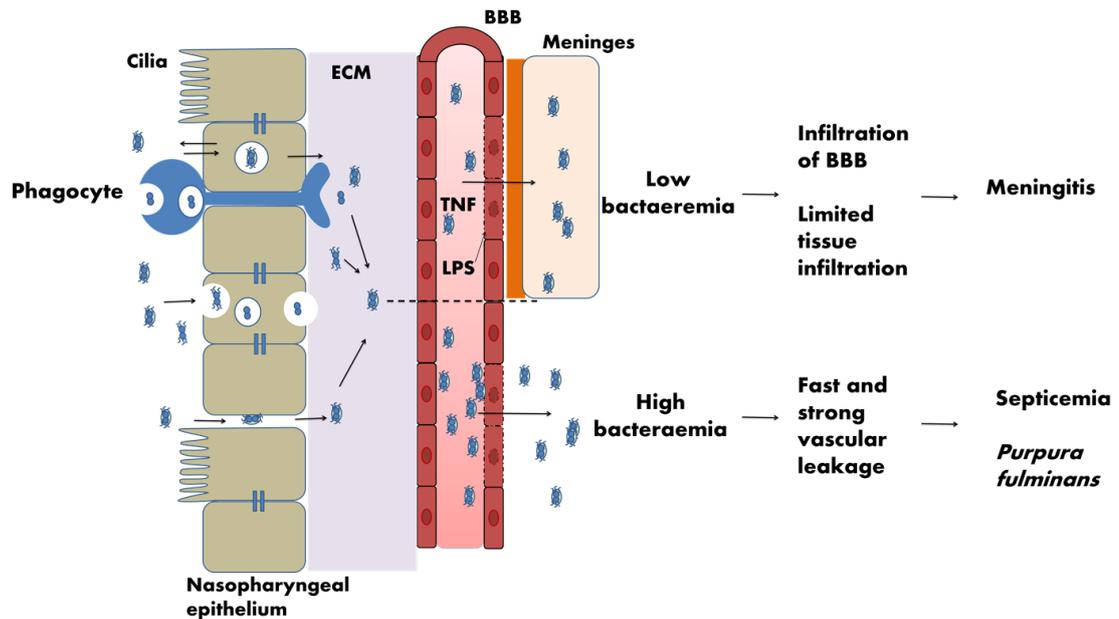
Figure 1.1 Schematic representation of the neisserial pathogenesis, with two different clinical outcomes.

N. meningitidis first colonises epithelial cells and establishes intimate contact with the non-cilliated host cells. *N. meningitidis* can cross the epithelium by transcytosis, through the paracellular route or using the 'Trojan horse' strategy – by using migrating phagocytes. Once in the blood stream, in susceptible hosts, *N. meningitidis* can multiply rapidly. Only a couple of neisserial microcolonies attached to the endothelium are sufficient to invade the blood-brain barrier (BBB) and rapidly disseminate in the cerebrospinal fluid, causing meningitis. Rapid and uncontrolled dissemination of the meningococci in the blood often results in septicaemia. This can progress within hours to the most serious implication of meningococcal disease, *purpura fulminans*, which has a mortality rate of over 50%. Picture adapted from Virji with permission (Virji, 2009), with modifications based on Coureuil *et al.* (Coureuil *et al.*, 2013).

Each stage requires a whole range of virulence factors to evade the human immune system and adapt to the challenging conditions during the course of

systemic infection. *Neisseria meningitidis'* ability to quickly adapt to the different hostile niches within the host organisms can be contributed to the flexibility of the meningococcal genome (Snyder *et al.*, 2005). The small size of the genome (2.2 Mbp), compared to other extracellular pathogenic bacteria, facilitates quick replication of the genome. This feature is advantageous for fast colonisation of the host mucosal surfaces inhabited with other competing commensal microorganisms (Lappann *et al.*, 2006). As a consequence, a fast replication rate results in an increased incidence of mutation, especially in a genome with an abundance of repetitive fragments. In *Neisseria* spp., this constitute 20% of the chromosomal DNA (Achaz *et al.*, 2002). Among the most important examples are short tandem sequence repeats termed 'simple sequence contingency loci' (SSCL). They are responsible for phase variability of the expression of many neisserial genes through slipped strand mispairing during replication (Lovett, 2004). Many neisserial genes linked to virulence such as PilIV, PorA, Opc, NadA and iron uptake proteins undergo phase variation. *Neisseria* spp*.* also increases their genetic variability by horizontal gene transfer. This is possible due to the existence of an extremely efficient natural transformation system, aided by a large number of DNA uptake sequences (DUS) located throughout the neisserial genome (Hamilton and Dillard, 2006). In addition, the presence of a large number of chromosomally integrated prophages, islands of horizontal gene transfer and insertion sequences with frequent homologous recombination and genome rearrangement all contribute to the high genomic diversity of *Neisseria* spp*.* (Schoen *et al.*, 2007). Such genome fluidity allows *Neisseria* spp. to develop a large repertoire of defence strategies to successfully colonise and survive within the host.

The mechanism of neisserial avoidance of the host defence mechanisms will be described in section 1.2.6 with the focus on the iron acquisition system as an example of withstanding the host nutritional immunity. A summary of host defence strategies with the neisserial major virulence factors involved in their evasion are shown in Figure 1.2.
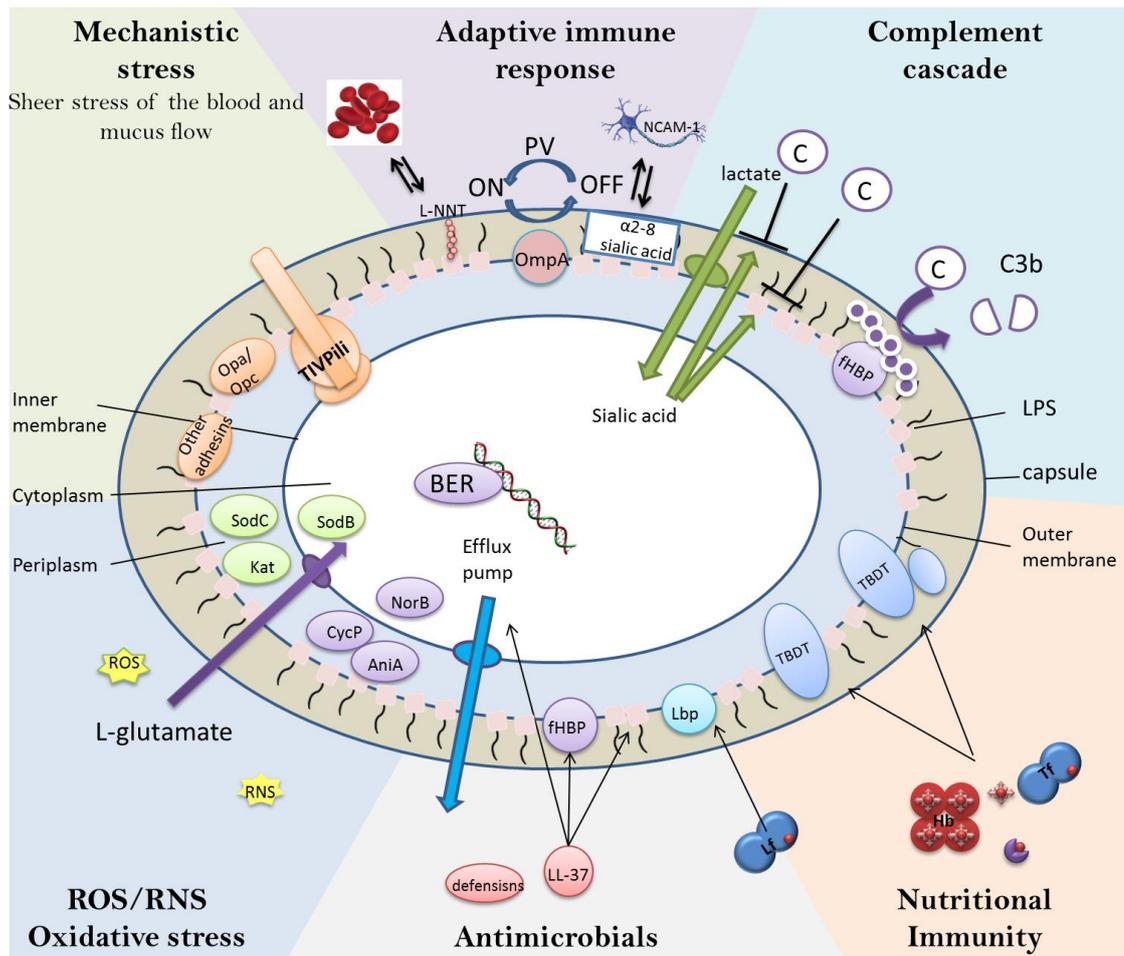
Figure 1.2 Summary of different host defence mechanisms against meningococcal infection with neisserial major virulence factors (MVF).

Mechanistic stress: Type IV pili is the MVF involved in the initial stages of colonisation of the epithelium and endothelium involved in the cellular signalling cascade. It promotes the formation of cortical plaques and microvilli, which protect bacteria from mucus and blood flow, and allow firmer attachment via other adhesins. Oxidative stress: Reactive oxygen and nitrate species (ROS and RNS respectively) are neutralised by a series of periplasmic and cytoplasmic enzymes. L-glutamate is converted to glutathione to keep the reducing environment in the cytoplasm. DNA oxidative damage is repaired with the Base Excision Repair (BER) pathway. Antimicrobials: Meningococci resist the antimicrobial activity of LL-37 thanks to the presence of LPS, binding of LL-37 to factor H binding protein (fHBP), and active transport of LL-37 via Mtr efflux pump. Lactoferrin binding protein (Lbp) protects meningococci from antimicrobial activity of lactoferrin (Lf) and simultaneously uses the Fe-complex as an iron source. Nutritional immunity: Host iron and haem sequestering proteins (haemoglobin − Hb, Transferrin (Tf) and Lf) and siderophores, produced by co-colonising bacteria, are used as an iron source via outer membrane single- or bi-partite TonB-dependent transporters. Complement cascade: The capsule and LPS prevent binding of the complement proteins. Sialic acid, which plays a major role in complement evasion, is produced from lactate which is actively transported inside the cells. Many pathogen proteins bind complement regulators, like fHbp binds factor H. Adaptive immunity: The presence of the α2-8 sialic acid homopolymer, which is identical to the NCAM-1 adhesin in host

neural cells, and the presence of L-NNT epitope in LPS, which is identical to the antigen in red blood cells, prevents recognition of the meningococci by the immune system. Phase variation (PV) of the major surface virulence proteins and their genetic variability allows evasion of the adaptive immune response and hinders vaccine development. Figure adapted from Lo *et al.* with modifications (Lo *et al.*, 2009). Permission of the publisher was granted. SOD – superoxide dismutase; Kat – catalase; AniA- nitrite reductase; NorB – nitric oxide reductase, CycP – cytochrome c'.

1.2.1 Adhesion strategies as an adaptation to the mechanistic stress of fluid flow in the mucosal surface and blood stream

During colonisation of the nasopharyngeal epithelium, a constant flow of mucosal fluid, together with the ciliary activity of the epithelial cells and other defence mechanisms such as coughing, impose a challenge for the successful attachment of pathogen cells and maintaining contact with epithelial cells. Blood borne meningococci must also overcome sheer stress of the blood flow in order to adhere to the vascular endothelium and progress to the meninges. Meningococci have evolved a whole array of outer membrane proteins and structures involved in the interaction with host cell receptors, allowing close adhesion to the host cells. These features give *Neisseria* spp. protection from mucosilliary clearance during colonisation, and allow successful colonisation of the endothelium and progression through the blood brain barrier.

1.2.1.1 Type IV pili mediated adhesion

Type IV pili are the main adhesins involved in the initial attachment of *Neisseria meningitidis* to the epithelial cells. Their filamentous structure can extend from the bacterial surface, which is of a particular importance for encapsulated meningococci with limited surface exposure of other adhesins. Pili have very dynamic structures. The constant process of pili extension and retraction, facilitated by the coordinated action of PilC and the ATP-ase PilT, is the basis of bacterial twitching motility (Merz *et al.*, 2000). In addition, pili are associated with competence for DNA transformation (Proft and Baker, 2009), formation of microcolonies and signalling to host endothelial cells (Lee *et al.*, 2005).
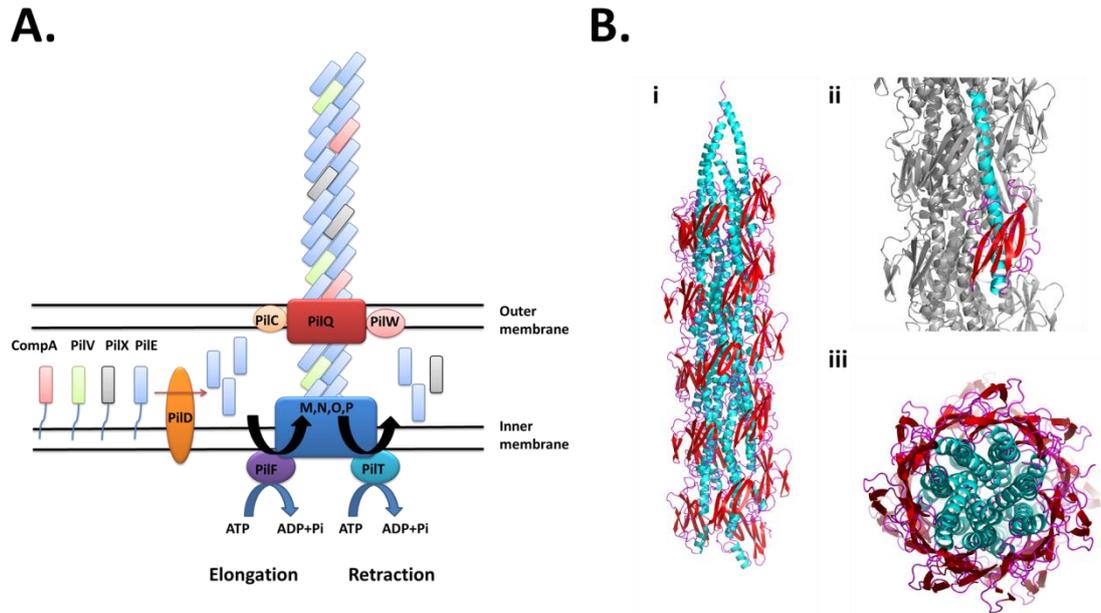
Figure 1.3 Type IV pilus assembly and structure.

A. Schematic representation of the type IV pilus assembly. The pilus fibre is comprised of the pilus subunits PilE, PilX, PilV and CompA, which are assembled by a platform complex (PilD, PilM-P,and PilF) located in the cytoplasmic membrane. They are transported to the cell surface through a pore formed by PilQ. The ATPase PilF delivers the energy for the elongation of the pilus, whereas PilT utilises energy from ATP for pilus retraction through the depolymerisation of the type IV pili. PilC (PilC1 and PilC2), located in the outer membrane, is involved in pilus retraction. In addition, PilC1 plays a role in adhesion to the host cells. PilW, located in the outer membrane, stabilises the PilQ structure. Figure adapted from Coureuil *et al.* with permission (Coureuil *et al.*, 2014). **B.** Ribbon representation of the type IV pilus fibre of *N. gonorrhoeae.* i. Structure of the pilus fibre. Pilin subunits are arranged helically with the alpha-helical domain (cyan) in the center of the pilus and β-stranded domain (red) exposed to the surface of the pilus. ii. An enlarged single pilin subunit. iii. A cross-section of the pilus fibre showing the arrangement of the alpha-helices inside the fibre, and the externally exposed β-strand domains. Images generated and rendered using Pymol from the PDB file 2HIL.

The fibre of the pilus is mainly composed of the phase variable major pilin, PilE, the subunits of which are arranged in a helical orientation (Virji *et al.*, 1992a) (Fig. 1.3). Several posttranslational modifications, such as the glycosylation or phosphorylation of serine residues, were observed in PilE. These alter the pilus surface charge, affecting adhesion and the pili's immunogenic properties (Virji, 1997). Some minor pilins (PilX, PilV and ComP) are also incorporated in the pilus fibre and until recently were believed to directly mediate adhesion, auto-

aggregation and host membrane reorganisation (Brown *et al.*, 2010). Recently, the results of the study by Imhaus and Dumenil demonstrated that the modulation of the Type-IV pili-mediated functions in meningococci depends on the number of pilis rather than the composition of the minor pilins in the pilus fibre. The role of the minor pilis (PilV and PilX), which according to this study are mainly located in the periplasm, is attributed to the initiation of the pilus biosynthesis and regulation of the number of pilis (Imhaus and Duménil, 2014)

The PilC proteins are proposed to form a structural tip of the pilus and are essential for adhesion to epithelial cells and pilus assembly; however, their exact role remains enigmatic. The pore forming PilQ pilins are also thought to be involved in adhesion (Rudel *et al.*, 1995).

The type IV pilus mediates adhesion of the meningococci to the epithelial cells via a recently identified receptor, CD147, and allows localised proliferation of the bacteria (Bernard *et al.*, 2014). The attachment of the bacteria facilitates the formation of membrane structures known as 'cortical plaques'. These structures are formed via the local recruitment of ERM proteins (ezrin-radizin-moeisin protein family) and several transmembrane receptors: Erb2 tyrosine kinase, CD44 and ICAM receptors which are attachment sites for the polymerised actin filaments (Merz *et al.*, 1999; Merz and So, 1997). Localised cortical actin polymerisation is activated by Rho GTPase and the Src tyrosine kinase signalling cascade, and activation of β-arrestins through β2-adrenergic receptors. The signalling cascade leads to the formation of membrane protrusions that initially act as a shield for the attached meningococci against the sheer force of the mucosal or blood flow. Later, elongation of the microvilli allows bacterial engulfment and internalisation (Merz *et al.*, 1996). The cellular receptors and signalling pathways used by *Neisseria* spp. for attachment are also implicated in leukocyte adhesion and trans-epithelial migration. Thus, it is hypothesized that *Neisseria* interferes with leukocyte adhesion by hijacking the leukocyte's docking sites (Doulet *et al.*, 2006).

Close contact with the endothelial cells also opens the para-cellular route for neisserial translocation through the endothelium. Activation of the β-adrenergic receptor and β-arrestins has been shown to result in the recruitment of proteins

from junctional complexes located under the neisserial microcolonies, leading to increase leakage of the endothelium (Coureuil *et al.*, 2013).

1.2.1.2 Opa and Opc

Although type IV pili are the first means of bacterial adhesion to the host cells, a firmer and more stable attachment is necessary for epithelial barrier crossing. This is achieved by interaction with two other groups of major adhesins: the Opa and Opc proteins, and LPS. Opa proteins belong to a family of transmembrane proteins with eight β-strands and four extracellular loops (Malorny *et al.*, 1998). The surface exposed loops are immunogenic, but their high degree of variability limits their potential as a vaccine candidates. Encoded by four to five loci, the proteins are also subjected to phase variation and homologous recombination. This greatly increases the repertoire of Opa variants expressed by a single strain (Callaghan *et al.*, 2006). Opa proteins bind to carcinoembryonic antigen cell adhesion molecules (CEACAM-1), heparan-sulphate proteoglycans (HSPG), present on some epithelial cells, and integrins (Dehio *et al.*, 1998; van Putten and Paul, 1995; Virji *et al.*, 1996). Opc proteins are encoded by a single gene, *OpcA,* and form a β-barrel consisting of ten transmembrane β-sheets. Their adhesion to integrins in the endothelial cell membrane is mediated via serum vitronectin or fibronectin. It has been shown that Opc can mediate cell adhesion independently of pili and Opa proteins in unencapsulated meningococci (Virji, 2009; Virji *et al.*, 1993; Virji *et al.*, 1992b; Virji *et al.*, 1995).

1.2.2 Evasion of the complement system

The complement system is a key factor of the host innate immune system, and consists of approximately 35 soluble or membrane bound proteins. The complement cascade is involved in the recognition and clearance of invading pathogens either in concert with antibodies and phagocytes, or independently (Morgan *et al.*, 2005). Three different pathways can activate the complement cascade: the classical pathway (initiated by the antibody-antigen interactions), the lectin pathway (activated by the interaction of mannose-binding-lectin with polysaccharide structures on the surface of the bacteria) or the alternative

pathway activated by spontaneous hydrolysis of C3. In all three pathways, the C3b fragment is released and causes the insertion of the membrane attack complex (MAC), increased opsonophagocytosis of invading pathogens, modulation of the inflammatory response and chemotaxis (Walport, 2001). Excessive activation of the complement cascade can be detrimental to the host, thus the activation process is tightly controlled. Several fluid-phase and soluble proteins are involved in the modulation of complement function. The negative circulating complement regulators include C4b-binding protein, factor H, vitronectin and C1q inhibitor (Lewis and Ram, 2014). Also, negative control is achieved by the expression of membrane bound complement receptors with inhibitory functions, such as Membrane Cofactor Protein (MCP) also known as CD46, complement receptor 1 (CR1), CD55 and CD59 (Kim and Song, 2006). Components of the complement cascade are mainly produced in the liver and circulated in the serum; although epithelial cells have been shown to express the complement proteins to protect mucosal surfaces of the nasopharynx and urogenital tract from pathogen invasion (Lo *et al.*, 2009).

Complement deficiencies, whether inherited or acquired, are important predisposing factors to meningococcal disease. This is especially evident in deficiencies of the complement components activated later in the cascade. For example, deficiencies in the membrane attack complex proteins (C5-C9) increase the risk of meningococcal disease by 7000 to 10000 fold when compared to the normal population (Figueroa and Densen, 1991). Complement deficiencies influence susceptibility to meningococcemia, severity and recurrence of the disease and mortality rates. Examples of the correlation between complement deficiencies with meningococcal disease progression or predisposition include: properdin (factor P) deficiencies, which are linked to severe fulminant septicaemia and high mortality rate (Braconier *et al.*, 1983), the mannose-binding-lectin (MBL) deficiency in children, which is associated with higher bacterial load during meningococcemia (Darton *et al.*, 2014), or factor H deficiency, which increases susceptibility to serogroup C *(Haralambous et al., 2006).*

Meningococci have developed a range of strategies that allow it to evade the human complement.

The capsule is the major determinant of the meningococcal serum resistance and virulence. Encapsulated meningococci are more resistant to complement mediated killing compared to strains lacking the capsule. Essentially all strains isolated from the blood are encapsulated, whereas only 50% of carriage isolates expressed a functional capsule (Lewis and Ram, 2014). Capsular polysaccharides of the serogroups B, C, W-135 and Y contain sialic acid which is also commonly present on the host cells. This makes them less susceptible to recognition by the immune system (Walport, 2001). The most effective mimicry strategy was developed by serotype B, where the structure of sialic acid residues in the capsule is identical to the one present on the neural cell-adhesion molecule (NCAM) on host cells. This particular feature prevented the use of this capsule polysaccharide in the composition of a vaccine against serogroup B, due to low immunogenicity and the risk of eliciting an autoimmune response (Diaz Romero and Outschoorn, 1994). On the host cells, sialic acid plays a role in down-regulating the alternative pathway through the enhancement of the interaction between factor H and C3b (Walport, 2001). This mechanism has not been proven for meningococci. The most probable function of the capsule in complement evasion is hindering the insertion of the membrane attack complex by masking the epitopes for the bactericidal antibodies (Drogari-Apiranthitou *et al.*, 2002; Lewis and Ram, 2014).

1.2.2.1 Lipopolysaccharide (LPS)

Lipopolysaccharide is a primary component of the outer membrane of all Gram-negative bacteria. It is responsible for outer membrane stability. In *Neisseria* spp*.* LPS expression is indispensable for the complement resistance (Geoffroy *et al.*, 2003). The structure of neisserial LPS consists of an inner and an outer core. The inner core contains several heptose molecules and is connected to the membrane anchored lipid A via one of the two KDOs (2-keto-3-deoxy-D-*manno*-2-octulosonic acids). The outer core is highly heterogeneous and consists of up to ten sugar residues. The short outer core of LPS differentiate *Neisseria* spp*.* from other Gram-negative bacteria, which usually express a longer side chain (up to 40 residues) consisting of repetitive sequences of sugar residues referred to as "O-antigen". Therefore, neisserial LPS is also known as lipooligosaccharide (LOS) (Jennings *et al.,* 1999). The composition of the LOS

side chain is the basis for further classification of *N. meningitidis* into twelve immunotypes (L1-L2) (Scholten *et al.*, 1994). In addition, modification of the sugar residues, such as the addition of sialic acid to the terminal lacto-*N*-neotetraose (LNT) epitope, increases the repertoire of LPS types, modulating the effectiveness in complement evasion (Mandrell *et al.*, 1991).

LPS modulates the complement cascade by the deposition of C3b and C4b (Lewis and Ram, 2014). Furthermore, sialylation of LPS enhances complement resistance in *Neisseria* spp*.* In *N. gonorrhoeae,* it hinders antibody binding to the neisserial surface (Elkins *et al.*, 1992), and enhances the binding of factor H to Porin B (Ram *et al.*, 1998). In *N. meningitidis* the exact role of LPS sialylation remains uncharacterised, although it appears to be involved in the modulation of the lectin pathway (Jack *et al.*, 1998; Jack *et al.*, 2001).

1.2.2.2 Binding of the host complement regulatory protein

In order to circumvent the complement killing machinery, *Neisseria* spp., like several other pathogens, express a number of proteins involved in hijacking the host complement regulatory proteins.

The negative regulator of the classical pathway, C4 binding protein (C4BP), is one of the neisserial targets for inhibition of the classical pathway. Strains of *N. meningitidis* that are able to bind C4BP were more resistant to complement killing than non-binders. PorA was identified as a binding partner of C4BP, however the interactions were observed under non-physiological conditions, undermining the importance of PorA in the modulation of the classical pathway (Jarva *et al.*, 2005).

Down-regulating the alternative pathway by Factor H binding is better characterised. Factor H deposition on *Neisseria* spp*.* results in the decrease of C3b deposition, thus increasing serum resistance of this pathogen (Lewis and Ram, 2014). *Neisseria* spp*.* expresses factor H binding protein (fHBP), also known as GNA1870, which is regarded as the main virulence factor of meningococci. Its selective binding to human factor H is one of the reasons for the narrow host species specificity of *Neisseria* spp*.* (Madico *et al.*, 2006). Factor H binding protein is a component of the Bexsero vaccine against serogroup B. Also, neisserial NspA, and recentlyPorB2, have been identified as

inhibitors of the alternative pathway, with an affinity for factor H (Lewis *et al.*, 2010; Lewis *et al.*, 2013). Presence of other factor H binding structures in *Neisseria* spp contributes to the limited strain coverage of the Bexsero vaccine. Widespread use of the vaccine is predicted to cause positive selection of the disease strains lacking fHBP isolated form patients (Giuntini *et al.*, 2013), thus prompting the need for a new generation vaccines against serogroup B.

Neisseria spp*.* also hijacks regulatory proteins at different levels of the complement cascade. The members of the opacity protein family: Opa and Opc have been shown to bind vitronectin to prevent the formation of the membrane attack complex (Blom *et al.*, 2009). Properdin and membrane-associated-complement receptor binding are also involved in the increase of serum resistance by bacteria (Lewis and Ram, 2014).

1.2.3 Antimicrobial peptides

Antimicrobial peptides and proteins are the first series of host innate defence mechanisms against pathogens. The antimicrobial peptides are small, cationic molecules that are either secreted by epithelial and immune cells into mucosal surfaces, or are present in neutrophil granules. They act as innate, broad spectrum antibiotics by promoting both the aggregation of bacteria and viruses, and their killing. Their role extends to the modulation of the immune response by promoting opsonisation and chemotaxis of neutrophils and monocytes, and by down-regulating inflammation (Seiler *et al.*, 2014). The antimicrobial peptides are divided into two groups: defensins (devided into three groups: α, β and θ) and cathelicidin. In the human host, only one cathelicidin has been described: LL-37 (Puklo *et al.*, 2008). *Neisseria* spp*.* are also subjected to other antimicrobial molecules at different stages of infection, such as elastase, lysozyme and lactoferrin (Lo *et al.*, 2009).

Neisseria spp*.* are largely resistant to the antimicrobial molecules produced by the host because they express a number of surface exposed structures. One example is the lactoferrin receptor LbpAB, which allows the utilisation of lactoferrin as an iron source, and will be further discussed in more detail in section 1.2.6.3.

The resistance to LL-37 and defensins is contributed to by the presence of the capsule, lipid A modifications in LPS, and biofilm formation by *Neisseria* spp*.* by limiting the access of the cationic peptides to the outer membrane (Jones *et al.*, 2009). Expression of the factor H binding protein also contributes to the neutralisation of the LL-37 anti-microbial effect, probably due to the electrostatic interaction between these two molecules (Seib *et al.*, 2009). Furthermore, the study by Tzeng *et al.* revealed that mutations in the operon encoding the efflux pump genes MtrC-MtrD-MtrE are associated with higher susceptibility to killing by LL-37 and other cationic antimicrobials, specifically polymyxin B and protegrin-1 (Tzeng *et al.*, 2005). This indicates that the efflux pump is important for sustaining antimicrobial killing mechanisms. It has also been shown that *Neisseria* spp*.* withstands the AMP challenge more effectively when attached to nasopharyngeal epithelial cells. The activation of the CDC42 and Rho cell signalling cascade in the host by meningococci, by a cytoskeleton-independent process, is believed to have a protective effect, although the mechanism is unclear (Geörg *et al.*, 2013).

1.2.4 Evasion of the oxidative stress

Role of professional phagocytes in generating oxidative stress

The mucosal surfaces, epithelium and blood stream contain an abundance of professional phagocytes that are involved in the clearance of the invading pathogens. Macrophages, monocytes and neutrophils differentiate from myeloid progenitor cells through a maturation process. Macrophages kill meningococci after activation through the classical complement cascade. As described before, meningococci are very proficient in circumventing the recognition and activation of the killing cascade. Neutrophils are the major cellular components of the immune system responsible for clearance of meningococci. Neutrophils use three characteristic killing strategies: 1. Intracellular killing *via* phagocytosis; 2. extracellular killing by the release of granules containing antimicrobial peptides and reactive oxygen and nitrite species that cause oxidative stress; 3. secretion of neutrophil extracellular traps (NETs) that are formed by secreted DNA, AMPs and ROS/RNS. The role of these NETs is to trap the pathogen and deprive it of essential nutrients and kill it (Criss and Seifert, 2012).

Neutrophils and macrophages both produce ROS (Reactive Oxygen Species) and RNS (Reactive Nitrogen Species) to destroy invading microorganisms. ROS/RNS target the bacteria in the phagolysosomes, and in the extracellular environment when released with the bactericidal granules. The enzyme responsible for the production of ROS is NADPH oxidase. This multiunit enzyme assembles on the phagocyte's plasma membrane in response to phagocyte activation, where it generates the production of hydroxyl free radicals, superoxide anions, and hydrogen peroxide (Roos *et al.*, 2003). In addition, NADPH oxidase is present in the granules of neutrophils alongside myeloperoxidase, which uses hydrogen peroxide to produce hypochlorous acid. This adds to the broad anti-microbial repertoire of neutrophil granules (Fang, 2004). RNS include nitric oxide produced by nitric oxide synthase (iNOS) and peroxynitrate. Apart from professional phagocytes, RNS is constitutively produced by epithelial cells of the respiratory tract as a cellular or intracellular signalling molecule. RNS regulate smooth muscle tone, ciliary function, electrolyte transport and protection from invading pathogens (Bove and van der Vliet, 2006).

Effect of ROS and RNS on pathogens

ROS/RNS are highly reactive and can penetrate the microbial envelope through anion channels (Pacher *et al.*, 2007). ROS/RNS react with thiols, metalloproteins and DNA. Oxidative damage of DNA is caused mainly by ROS, whereas RNS interferes with respiratory processes and DNA replication by oxidising metal ions (Fe, Mn) required by key enzymes involved in these processes (Fang, 2004). Bacteria appear to have evolved many mechanisms to evade killing by ROS and NOS. These are considered below.

Neutralisation of ROS by meningococci

The first line of defence from ROS in meningococci is the enzymatic neutralisation of ROS by superoxide dismutases and catalases (Criss and Seifert, 2012). Highly unstable superoxide anions are converted to hydrogen peroxide by superoxide dismutase. The catalases neutralise hydrogen peroxide to producing water and oxygen (Eason and Fan, 2014). Meningococci express two types of superoxide dismutases: cytopolasmic SodB and periplasmic SodC (Wilks *et al.*, 1998). SodC is also involved in the neutralisation of ROS released

by activated macrophages, and protects opsonised meningococci from phagocytosis (Dunn *et al.*, 2003). Another mechanism of protection from oxidative stress is the uptake of L-glutamate via the L-glutamate transporter (GltT), and converting it into glutathione. As a result, the redox environment in the cytoplasm of the pathogen helps to reduce the damage by ROS (Talà *et al.*, 2011).

Neutralisation of RNS by meningococci

Meningocci are adapted to microaerobic growth by utilising nitrate produced by two enzymes from the denitrification pathway: nitrate reductase (AniA), and nitric oxide reductase (NorB). AniA catalyses the conversion of nitrate into nitric oxide, whilst NorB reduces NO to the more stable nitrous oxide (Rock and Moir, 2005). These enzymes help *Neisseria* spp. to detoxify RNS with help from cytochrome C', the product of the *cycP* gene (Anjum *et al.*, 2002). Expression of NorB and cytochrome C' has been shown to increase the survival of meningococci in the nasopharyngeal mucosa and in macrophages by different mechanisms (Stevanin *et al.*, 2005). Depletion of NO, which is an important signalling molecule, has been attributed to the modulation of the inflammatory response exerted by monocyte derived macrophages by changing the levels of cytokines and chemokines (Stevanin *et al.*, 2007). It also promotes intracellular survival of phagocytised meningococci by inhibiting macrophage apoptosis (Stevanin *et al.*, 2005). According to the study by Laver *et al.*, this phenomenon can be explained by the reduction in NO, which decreases the level of S-nitrosylation of the host cells (Laver *et al.*, 2010). S-nitrosylation is an important post-translational modification of proteins, and is implicated in the control of intracellular cell signalling, apoptosis, neurotransmission and gene expression. Therefore, by metabolising NO, the main substrate for S-nitrosylation, meningococci are able to modulate many cellular and inflammatory processes in the host for its own advantage.

DNA repair

Damage of DNA caused by oxidative stress can, if unrepaired, lead to abnormal cell function, genome instability during replication, and accumulation of mutations which can affect pathogen survival (Friedberg EC, 2006). The primary mechanism of DNA repair after oxidative stress in cells is the base

excision pathway (BER) (Bjelland and Seeberg, 2003). Meningococci have a lower number of BER enzymes compared with *E. coli,* and do not possess an SOS repair system (Black *et al.*, 1998; Davidsen and Tønjum, 2006). However, their repair network is efficient enough for withstanding oxidative damage. Two key enzymes of the BER system in meningococci are the bi-functional glycosylases Nth and MutM (also known as Fpg), which are involved in the removal of damaged nucleobases and incision of the DNA-backbone by two different mechanisms (Nagorska *et al.*, 2012). In addition, MutY, a mono-functional glycosylase, has an anti-mutagenic role by removing adenines paired with 8-oxoG, a common product of oxidative damage (Davidsen *et al.*, 2005).

The abasic nucleotides are further fixed by AP endonucleases or exonucleases by BER (Carpenter *et al.*, 2007; Davidsen *et al.*, 2007). Deletion of both Nth and MutM genes greatly decreases the survival rate of *N. meningitidis* that are subjected to oxidative stress (Nagorska *et al.*, 2012).

1.2.5 Evasion of the adaptive immune system

Most people are likely to be a carrier of *Neisseria meningitidis* at some stage of their lives, but only a small proportion of individuals develop meningococcal disease. Protection against *N. meningitidis* is exerted by the presence of the anti-meningococcal antibodies in healthy patients, which is quite common in the general population (Goldschneider *et al.*, 1969). Anti-meningococcal antibodies can clear the pathogen by three different mechanisms: 1. Bactericidal mechanism by the activation of the classical complement pathway and the insertion of the membrane attack complex (Goldschneider *et al.*, 1969); 2. Opsonophagocytosis, where antibodies bound to the pathogen surface promote its uptake by neutrophils (DeVoe, 1976); 3. Antibody-dependent cellular toxicity in sites with low levels of complement (Lowell *et al.*, 1979). Immunity against meningococcal infection can be acquired naturally by exposure to non-neisserial organisms (i.e. *E. coli*), through carriage of commensal *N. lactamica* or asymptomatic carriage and disease associated with *N. meningitidis.* Carriage of *N. lactamica* has been shown to generate cross-reactive antibodies that are protective against *N. meningitidis* colonisation. However, they have been shown to promote opsonophagocytosis rather than a bactericidal effect (Evans *et al.*,

2011). Carriage of meningococci or past history of meningococcal disease elicit the production of anti-meningococcal IgA and IgG antibodies, which protect the host from both colonisation of the nasopharynx and invasion of the blood stream. Vaccination is another way of acquiring immunity against pathogenic *N. meningitidis.*

Although naturally acquired immunity seems to be quite efficient in the control of meningococcal invasion, there are still meningococcal strains which can evade recognition by the host's adaptive immune system. In most of the serogroups, vaccination has proved to be efficient in preventing infection. However, neisserial genomic fluidity and rapid genomic changes in concert with antigenic pressure activated through the vaccination process leads to the selection of strains which are no longer sensitive to the vaccine elicited bactericidal killing.

Antigenic variability of the major neisserial virulence factors is a huge problem in the development of cross-reactive neisserial vaccine. *Neisseria* spp*.* have the largest number of genes that undergo phase variation in all pathogenic bacteria (Saunders *et al.*, 2000). Phase variation, alongside other mechanisms promoting genomic rearrangements (which were described in Section 1.2), contributes to the huge repertoire of genes and rapid changes which lead to the loss of vaccine induced immunity.

Serogroup B uses another strategy to evade recognition by the immune response named molecular mimicry. The capsule and LPS of this serogroup is chemically modified so that it resembles structures present on the host cells. The capsule of serogroup B contains a homopolymer of α2-8 sialic acid, which is also present on the NCAM-1 antigen on host neural cells (Finne *et al.*, 1983). The LPS molecules contain a lacto-N-neotetraose epitope (L-NNT), which is also present on the surface of human red blood cells (Lo *et al.*, 2009; Moran *et al.*, 1996).

As a consequence, the capsule and LPS of serogroup B are characterised by low immunogenicity. In addition, their inclusion in the vaccine formulation could result in eliciting auto-immune antibodies.

1.2.6 Defence against host nutritional immunity - iron acquisition by *Neisseria* spp.

Iron is an essential nutrient for almost all living organisms. To date, the few exceptions reported are microorganisms such as *Lactobacillum plantarum* and *Borrelia burgdorferi* who use manganese or cobalt in place of iron (Posey and Gherardini, 2000; Weinberg, 1997). Iron is available in two oxidation forms: as ferrous (Fe^{2+}) and ferric (Fe^{3+}) ions. Changes in the redox state of iron are utilised by many enzymes involved in cellular functions. However, by-products of cellular oxidative reactions in the presence of iron are converted into reactive oxygen species, which are highly toxic to the cells. Free radicals can damage lipids, DNA and proteins. Thus, iron levels, and its subcellular localisation, are tightly regulated.

In the human host, the majority of free iron is stored within cells as haem, iron-sulphur clusters in various proteins or in ferritin, which contains 15-20% of total body iron (Clarke *et al.*, 2001; Jordan and Saunders, 2009). The main source of intracellular haem is haemoglobin (Hb) which constitutes 67% of the total haem, and is normally contained within erythrocytes (Runyen-Janecky, 2013). Other intracellular sources of haem are myoglobins, cytochromes and the recently described cytoglobin and neuroglobin (Liu *et al.*, 2012b; Watanabe *et al.*, 2012). Extracellular iron in the host is usually found in complex with iron-binding proteins including transferrin (Tf) and lactoferrin (Lf). Each of the iron-binding proteins has a specific niche in the human body i.e. transferrin is predominantly found in serum, but is also present in CSF, whilst lactoferrin is in secreted fluids such as tears, milk and bile (Jordan and Saunders, 2009). Spontaneous haemolysis causes the release of Hb in the serum, but free Hb is captured by the serum haptoglobin (Jordan and Saunders, 2009). Free haem in the serum released by turnover of haem-binding proteins is captured by haemopexin, which has the highest known affinity to haem (Tolosano and Altruda, 2002), with a K_d in the subpicomolar range (Hrkal *et al.*, 1974).

Sequestration of iron and haem within host proteins not only protects the organism from the toxic effects of these molecules, but it is also believed to play an important role in non-specific protection from infections known as "nutritional

immunity" (Weinberg, 1975). Iron availability is an important limiting factor for neisserial invasion, as free iron concentration in the host $(10^{-9} – 10^{-18}$ M) 10^{12}-fold lower than that required for microbial growth (Cassat and Skaar, 2013; Weinberg, 1978). Levels of free iron in the host can be lowered even more dramatically in response to infection by the release of the liver hormone hepcidin, which is responsible for iron homeostasis (De Domenico et al., 2010; Rossi, 2005). Hepcidin is also released locally at the site of infection by neutrophils and macrophages, thus limiting iron availability (Peyssonnaux et al., 2006). Other inflammatory cytokines, such as interferon-gamma, TNF-alpha, IL-1 and IL-6, can also induce hypoferremia in a hepcidin-independent manner by modulating local iron metabolism (Nairz et al., 2010; Weiss, 2005). In addition to these mechanisms, neutrophils at the site of infection release secondary granules which contain lactoferrin. Lactoferrin has a higher affinity for iron compared to other iron binding proteins in the acidic environment characteristic of sites of infection, and can contribute to local hyperferremia (Baker and Baker, 2012). Another mechanism of iron sequestration is the secretion of lipocalin-2, also known as siderocalin or neutrophil gelatinase-associated protein (NGAP), by neutrophils, macrophages and other cell types in response to the inflammation. Lipocalin-2 can bind directly to certain siderophores, like enterobactin (Bachman et al., 2009; Smith and Wilks, 2012). Also, phagocytes mobilised at infection sites are believed to limit iron availability by overexpressing the NRAMP1 intracellular iron transporter. NRAMP1 is responsible for a reduction in the amount of divalent ions in the endosomes, which can have implications for the intracellular survival of pathogens (Jabado et al., 2000; Vidal et al., 1995).

Pathogens seem to have evolved a number of strategies to overcome host iron-sequestering defences. The mechanisms of iron acquisition depend on the pathogen niche (for example mucosal surfaces, blood stream or cerebrospinal fluid - CSF), localisation (intra- or extracellular)and preferred iron source (for example haemoglobin, transferrin, lactoferrin, haem). Pathogens modulate the expression of proteins involved in iron acquisition according to the availability and the source of iron. This review will cover iron acquisition strategies of gram

negative bacteria, with the focus being on the neisserial iron acquisition system (summarised in Figure 1.11).

1.2.6.1 Siderophores

Siderophores are low molecular weight molecules secreted by Gram-positive and Gram-negative bacteria to chelate free iron from precipitates or host iron-binding proteins in the extracellular milieu (Krewulak and Vogel, 2008). Although very diverse in structure, they can be divided into three groups based on the functional groups that are involved in iron coordination: hydroxycarboxylate, catecholate or hydroxymate (Boukhalfa and Crumbliss, 2002). Siderophores can bind iron tightly, with a dissociation constant that exceeds 10^{-50} M (Cassat and Skaar, 2013). This allows them to compete with transferrin. Ferric iron–siderophore complexes bind to designated outer membrane receptors, and are transferred inside cells by TonB-dependent transporters, which will be discussed in more detail in section 1.2.6.3. Once inside the cell, ferric iron is stripped from the siderophore, which is then either degraded, or recycled and secreted back into the extracellular environment. Although in *Neisseria* spp. no siderophores were yet identified, it has been shown that *N. gonorrhoeae* is able to bind siderophores produced by other pathogenic bacteria (Carson *et al.*, 1999).

1.2.6.2 Haemophores

Functions of haemophores are similar to siderophores. However, unlike siderophores, haemophores are proteinacious in nature and acquire Fe-haem complexes from host haem-transporter proteins. Hemophores are either secreted, or are loosely associated with the outer membrane. They act by either outcompeting the host proteins (like HasA from *Serratia mercescens*), or by inducing conformational changes in the host proteins, resulting in the release of haem from the complexes (for example HxuA from *H. influenzae*). Haemophores loaded with haem interact with the appropriate TonB-dependent outer membrane receptor and either release haem, which is transferred to the cytoplasm, or the haemophores remain associated with the outer membrane and function as haem storage proteins (Wandersman and Delepelaire, 2012). The best characterised haemophore expressed by Gram-negative bacteria is HasA from *Serratia marcescens*, which has sequence homologs in

Pseudomonas aeruginosa and *Yersinia pestis* (Létoffé *et al.*, 1998; Rossi *et al.*, 2001). HasA acquires haem from haemoglobin, haemopexin, leghaemoglobin and myoglobin (Wandersman and Delepelaire, 2012). Upon haem binding, HasA undergoes a conformational change. It has been shown that HasA can form dimers by domain swapping. The HasA dimers bind haem with lower affinity than the monomers. However, the dimers are not able to bind to HasR, indicating that this might be a mechanism of haem storage and sequestration to limit availability of this molecule from other, competing microorganisms (Czjzek *et al.*, 2007). The structure of the HasA-HasR-Haem complex is shown in Figure 1.4.

In many cases, haemophores aid haem acquisition from host binding proteins, but are not essential for the function of outer membrane receptors. In the neisserial HpuAB bipartial transporter, both components of the haemoglobin-haptoglobin receptor are indispensable for the extraction of haem from the host haem-complexes. HpuA is a lipoprotein anchored in the outer membrane, like many bacterial haemophores, however, when expressed alone, HpuA is unable to bind either haemoglobin or haemoglobin-haptoglobin complexes. The outer-membrane receptor HpuB, when expressed alone, showed detectable binding of the Hb and Hb-Hp complexes, but this phenotype was unable to utilise haemin from the bound complexes. This indicates that HpuA does not function as a haemophore *per se*, but is an essential component for substrate recognition and transport across the outer membrane (Rohde and Dyer, 2004). No other haemophore proteins have been identified in *Neisseria* spp. to date.
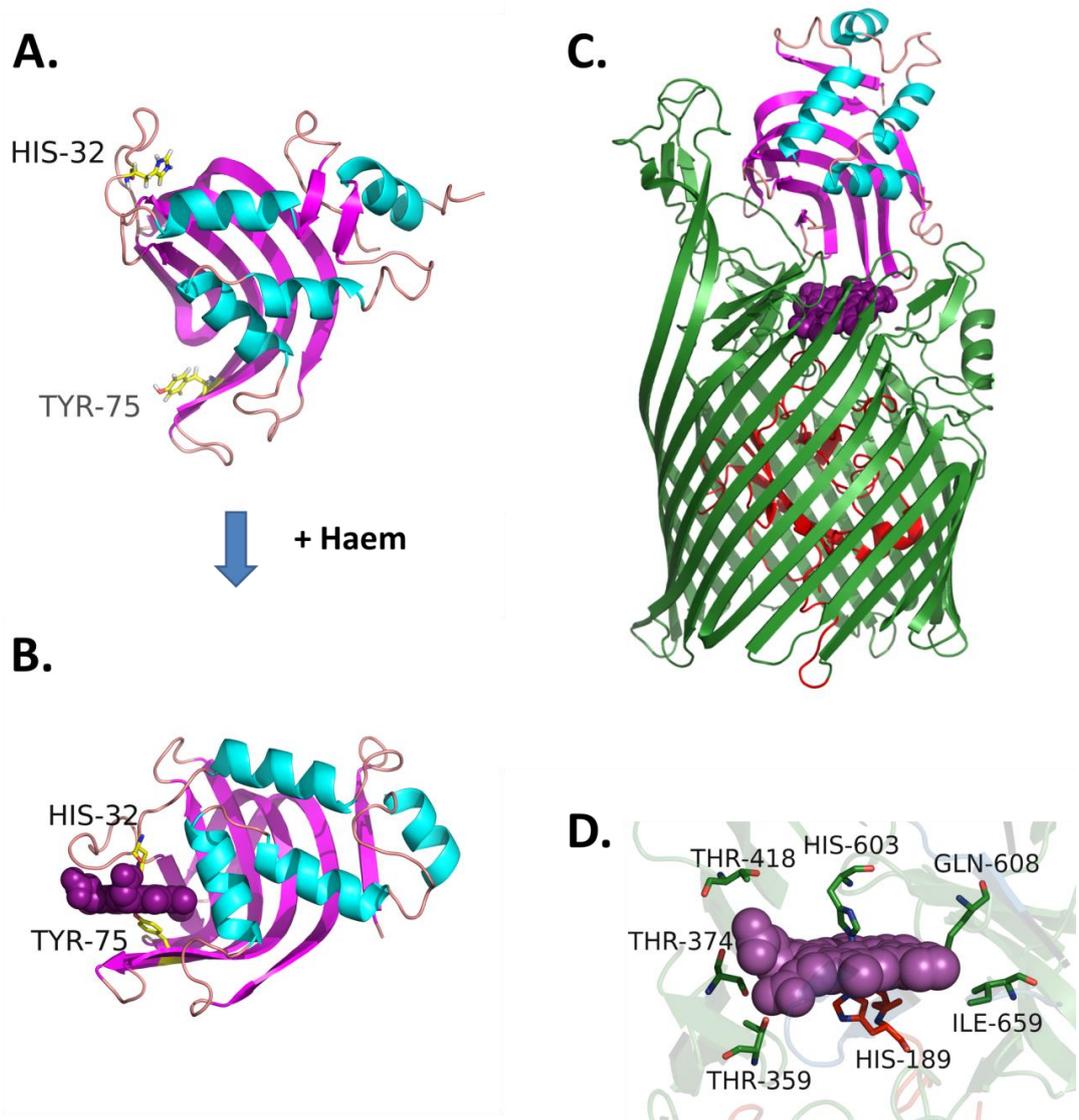
Figure 1.4 Representation of structures of haem receptor HasAR proteins of *Serratia marcescens*.

A. Ribbon representation of the structure of apo-HasA in an open conformation (PDB file: 1YBJ) **B.** Structure of holo-HasA (PDB file: 1B2V). The haem molecule is shown in purple, as a space-filling representation. Residues involved in haem binding (haem-32 and Tyr-75) are represented as yellow sticks. The loop containing His-32 closes on haem upon its binding. **C.** Structure of HasR, a TonB dependent haem transporter, in complex with haem and HasA (PDB file 3CSL). HasR forms a 22-β-strand barrel structure (shown in green) with the plug domain occluding the pore of the receptor (represented in red) – a structure typical for TonB-dependent receptors. Binding of holo-HasA to HasR triggers a conformation change, with the protein returning to its open conformation, which allows the release of haem. **D.** HasR residues within 5 Å of the bound haem molecule are shown in stick representation. His-189 is a residue in the loop of the plug domain, whereas His-603 is a highly conserved residue in many of the bacterial haem receptors (as shown in Fig. 1.8). Images were generated and rendered in Pymol.

1.2.6.3 TonB-dependent transporters (TBDTs)

The outer membrane of Gram-negative bacteria constitutes an impermeable barrier for siderophores, haem and iron complexes. Only free iron can utilise the outer membrane porins to cross this barrier. However, this transport system is extremely inefficient due to the low levels of iron in the extracellular milieu. Bacteria overcome this hurdle by expressing outer membrane receptors, which either directly transfer siderophores and haem across the membrane into periplasm, or are involved in the extraction of iron or haem from host transport proteins or from their own haemophores. The energy for transport is provided from the proton motive force of the inner membrane by the TonB/ExbB/ExbD complex. The TonB-dependent transporters can either bind directly to the ligand, extract iron or haem and transport it across the membrane (single component TBDTs) or, in addition to the membrane receptor, utilise another component in the form of lipoproteins in iron acquisition (bipartite TBDTs). FetA is an example of single component neisserial receptor for xenosiderophores.

Structure of TBDTs

In 2010 Noinaj and coworkers published a comprehensive structural analysis of the twelve solved TBDT structures available at that time (Noinaj *et al.*, 2010). Recently reported structures of transferrin binding proteins A and B with bound transferrin were analysed by negative electron stain microscopy. The analysis of tripartite structure consisting of TbpA, TbpB and human holo-transferrin greatly enhanced the understanding of structure and function of the iron outer membrane transporters in Gram-negative bacteria (Calmettes *et al.*, 2012; Noinaj *et al.*, 2012b).

All TBDT structures that have been solved share similar topology despite the limited sequence homology and different ligands. All these proteins consist of a 22-stranded β-strand barrel with an enclosed N-terminal plug-domain that occludes the pore to prevent the entry of substances harmful to the bacteria. In some structures, the plug domain is buried inside the β-barrel structure, whereas in others it protrudes into the periplasm. The N-terminus of the plug domain consists of a TonB box, which is a consensus sequence responsible for interaction with the TonB-ExbB-ExbD complex, and transduction of energy from the inner membrane. Plug domains are highly solvated indicating possible

conformational changes that allow the transport of ligands through the pore. Plug domain loops can protrude into the extracellular space and be involved in the binding of ferric or haem-complexes (Noinaj *et al.*, 2010).

The transmembrane β-sheets are connected to the periplasmic site by 10 short loops consisting of 2 to 10 residues. The extracellular receptor site consists of 11 long loops, labelled from L1 to L11, with sizes ranging from 2 to 37 residues. These extracellular loops are not only responsible for interacting with the ligand, but they can also form flaps occluding the light of the pore (Krewulak and Vogel, 2008; Noinaj *et al.*, 2010). Extracellular loops show highest sequence diversity even between close homologs. This will be as a result of evolutionary divergence in response to different ligands and antigenic variation.

Iron transporters – TbpAB and LbpAB

To utilise iron from the host proteins lactoferrin and transferrin, *Neisseria* spp. uses bipartite TonB-dependent transporters: LbpAB and TbpAB respectively. The transmembrane proteins LbpA and TbpA show a typical topology for TonB-dependent transporters. They share 43% sequence identity, with the highest level of diversity found in the external loops. Some unique features were identified in the TpbA structure, including the presence of a helix finger in loop 3 that is involved in iron extraction, extremely long loops that are involved in transferrin binding (loop 4 and 5), and a long loop in the plug domain that contains the conserved iron binding motif EIEYE. Additionally, loop 8 appears to be necessary for iron internalisation from transferrin (Noinaj *et al.*, 2012b) (Fig. 1.5. B).

Despite different ligands and low sequence identity (below 30%), the lipoproteins LbpB and TbpB do have similar structures (Fig 1.5 A.). The structures of these proteins are characterised by the presence of two domains known as lobes (N- and C-terminal lobe). Each lobe consists of a barrel formed by 8 β-strands, which is flanked by a handle domain rich in β-strands. Additionally, the N-terminal lobes contain an α-helical fragment. Both lobes are connected by a linker region, which is longer than the LbpB protein. The main feature that distinguishes between LbpB and TbpB is the presence of a negatively charged region in the C-lobe of LbpB, which is believed to be

involved in protection from the cationic antimicrobial peptide lactoferricin (Morgenthau *et al.*, 2012). Also, TbpB remains attached to the outer membrane via a C-terminal anchor peptide, which is required for interaction with TbpA (Yang *et al.*, 2011). In contrast LpbB is subjected to cleavage by the NalP membrane protease, and can be released into the extracellular environment. One hypothesis states that LbpB function can be modulated via NalP activity. In the blood, where NalP is overexpressed, more LbpB is released in response to a more hostile environment, whereas on mucosal surfaces, more LbpB remains surface associated due to the limited activity of NalP, and is involved in lactoferrin binding (Morgenthau *et al.*, 2014).
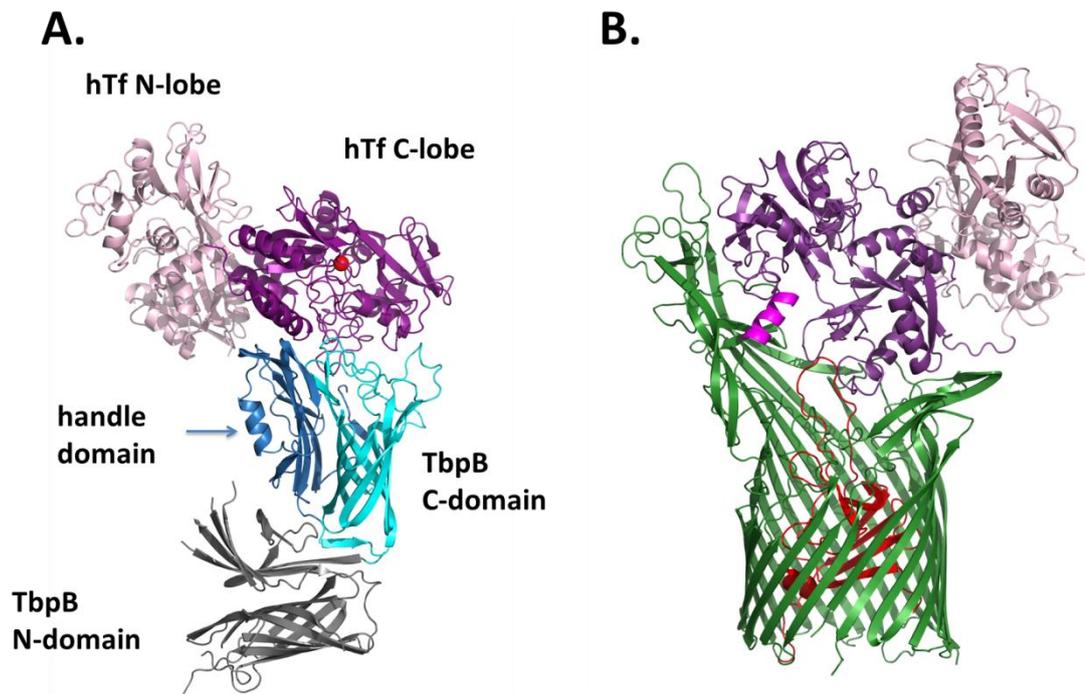
Figure 1.5 Crystal structures of TbpB and TbpA in complex with human transferrin.

A. A cartoon representation of TbpB in complex with holo-human transferrin with iron depicted as a red sphere. TbpB binds iron-loaded human transferrin and stabilises its closed conformation. **B.** A cartoon representation of TbpA in complex with human apo-transferrin. TbpB is a typical TonB-dependent transporter consisting of 22 anti-parallel β-strands that form a pore in the outer membrane (green) with a plug domain that occludes the light of the channel (red). TbpA consists of an L3 alpha-helix finger, which is involved in iron release from transferrin. Images were generated using Pymol using with PDB files: 3VE1 and 3V8X.

Neither lipoprotein is essential for the survival of bacteria when functional outer membrane receptors are present, but their presence increases the efficiency of iron acquisition. TbpB shows a preference for the iron-loaded forms of transferrin, whereas TbpA binds both forms, apo- and holo-Tf, with similar affinity. It has been speculated that surface exposure of the lipoprotein through the anchor peptide, and preference for the holo-form of Tf, allows more efficient and unrestricted capturing of the iron-loaded ligand in the environment with lower iron availability (Silva *et al.*, 2012). Other proposed functions of the lipoproteins include participation in iron extraction after binding to the outer membrane receptor, and facilitating the dissociation of the ligand from the membrane receptor after iron is removed (Noinaj *et al.*, 2012a).

Mechanism of iron utilisation from transferrin by TbpAB transporter

Analysis of recently solved crystal structures of transferrin binding proteins has allowed an increased understanding of the iron extraction process from the host proteins, and transport of the iron through the outer membrane receptor. The TbpB N-lobe has been shown to be a binding site for the C-lobe of holo-transferrin. The interactions between these two proteins cause conformational changes in the transferrin, resulting in the stabilisation of the iron-loaded form of the protein. The binding promotes the conformational changes in the anchor domain of TbpB, and allows a closer interaction with TbpA.

It is not exactly clear how the transferrin is passed from TbpB to TbpA, and whether it involves the simultaneous binding of transferrin by those two proteins. Two studies have produced contradictory models. According to Noinaj and colleagues (Noinaj *et al.*, 2012a) the transferrin has non-overlapping binding sites recognised by TbpA and TbpB, allowing formation of the transient tri-partite complex (Fig. 1.7), whereas another study claims that the binding sites overlap, preventing simultaneous binding by those two proteins (Silva *et al.*, 2012). No direct interaction between TpbA and TpbB has been observed when TbpB is in the apo-form. Upon binding to TbpA, transferrin changes its conformation to an open state, which allows iron extraction. This process is probably dependent on the energy derived from TonB. It has been shown that a single lysine residue in the helix finger of loop 3 plays an important role in this process. Furthermore, transport of iron through the TbpA pore is coordinated by

changes in the entire plug domain. Transferrin dissociates from TpbA, aided by TpbB, which shows a much lower affinity to apo-transferrin (Noinaj *et al.*, 2012b). A schematic representation of the iron acquisition process from human transferrin by the neisserial TbpAB transport system is shown in Figure 1.6.
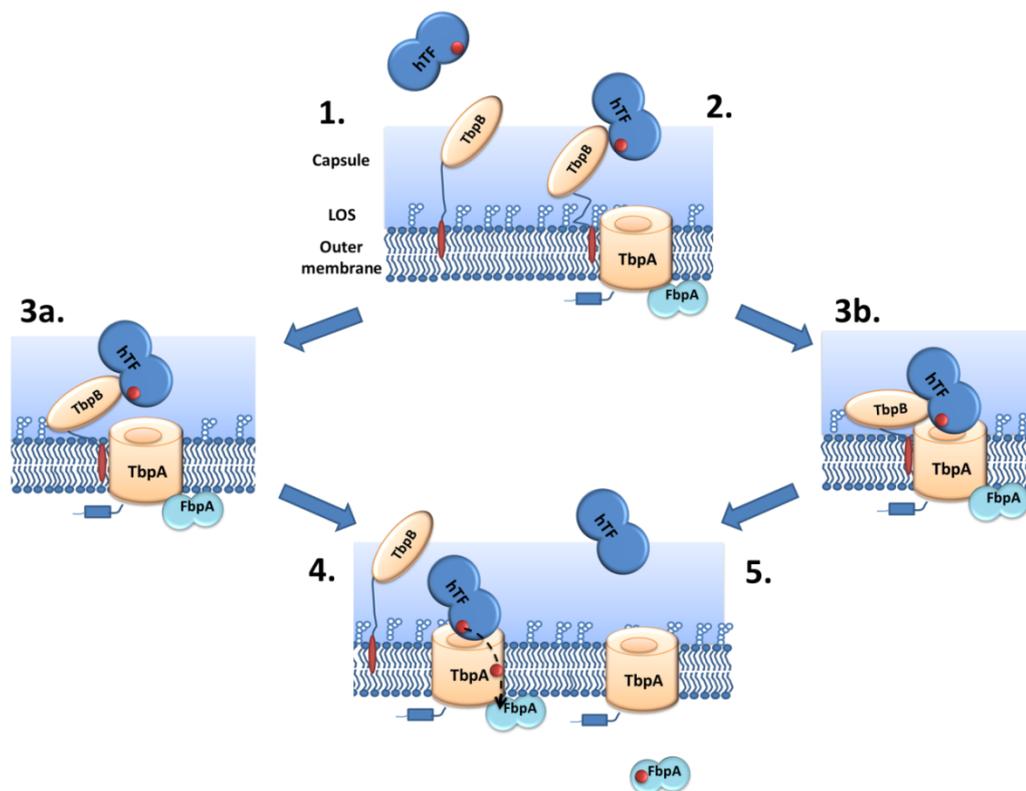
Figure 1.6 Model of iron transport by TbpAB receptor.

1. Holo-human transferrin (hTF) is captured by TbpB, with the docking peptide in the fully extended conformation. **2.** Upon binding of hTf, the docking peptide undergoes a confirmation change bringing hTF closer to the TbpA. **3.** The two models of transfer of hTF from TbpB to TbpA: **3a**. Interaction of TbpA with the anchor peptide of TbpB results in the release of hTF that can be readily captured by TbpA. Simultaneous binding of hTF by TbpA and TbpB is prevented by steric and allosteric factors. (Silva *et al.*, 2012). **3b.** TbpB transfers hTF to TbpA with the formation of a transient tri-partite complex, as proposed by Noinaj *et al.* (Noinaj *et al.*, 2012a). In both models, energy from the TonB system is required for the dissociation of TbpB from hTf. **4.** Iron is released from hTF and transported through the pore of TbpA in the periplasm, where it is captured by FbpA docked to the periplasmic loop of TbpA. **5.** Apo-hTf and holo-FbpA dissociate from TbpA. Figure based on (Noinaj *et al.*, 2012a; Silva *et al.*, 2012)
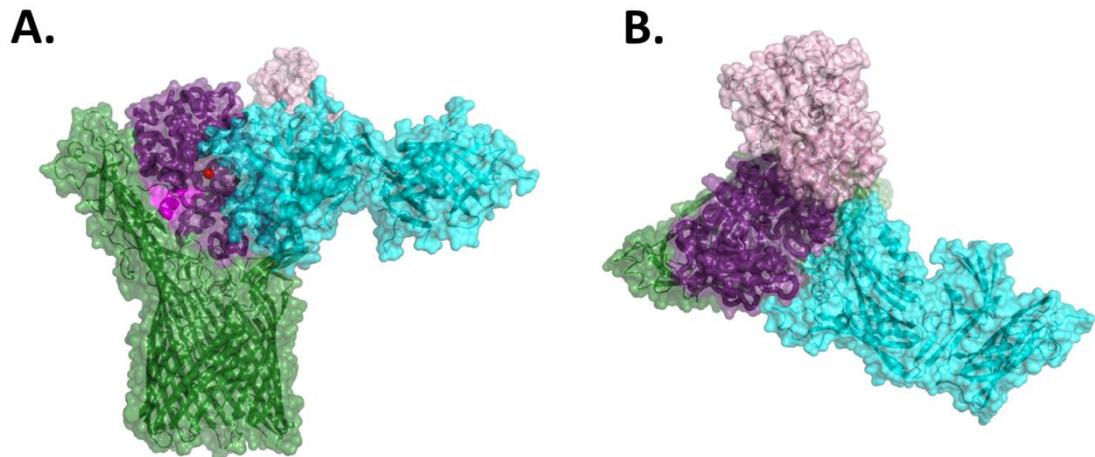
Figure 1.7 Model of the structure of the tri-partite complex of TbpA and TbpB with human transferrin.

A. A view from the site of the complex, showing simultaneous binding of the transferrin C-lobe by TbpA and TbpB. **B**. An overhead view from the top of the tri-partite complex. The model is based on the structures PDB ID: 3V8X and 3VE1, and is proposed by Noinaj and colleagues and supported by the data from the negative stain electron microscopy (Noinaj *et al.*, 2012b). Both panels were drawn and rendered in Pymol. The cartoon representation with the transparent surface shows TbpA coloured in green, with the L3-loop finger in magenta, TbpB is shown in cyan, the transferrin C-lobe in purple, and the N-lobe in pink. A ferric ion is shown as a red sphere.

Unlike HpuAB, transferrin and lactoferrin transporters in *Neisseria* spp. show species selectivity with regards to the binding ligand. This is because the region of transferrin that interacts with the TbpAB receptor is highly variable in other species (Noinaj *et al.*, 2012a).

The Role of the TonB-complex in iron transport

Transport through the outer membrane via TonB-dependent transporters requires energy. As the periplasm is devoid of ATP, and the outer membrane has no electrochemical potential to generate energy, TonB-dependent transporters utilise proton motive force from the inner membrane delivered by the TonB/ExbB/ExD complex. The TonB protein is anchored in the inner membrane via its N-terminal domain. It contains a proline-rich central domain which spans the periplasm. Its C-terminal domain is responsible for the

interaction with the outer membrane receptor by its TonB-box. ExbB and ExbD are localised in the inner membrane and support TonB. The ExbB protein has three transmembrane helical domains that interact with TonB. ExbD in turn stabilises the ExbB structure. ExbB and ExbD harvest energy from the transmembrane electrochemical potential and transfer it to TonB. Upon ligand binding, the outer membrane receptor undergoes a conformational change, allowing the interaction of TonB with the N-terminal TonB-box. The energy delivered by TonB causes unfolding, or pulling, of the plug domain of the membrane receptor, facilitating transport of the ligand by an unknown mechanism. In many cases, energy is also needed for dissociation of the apo-haemophore or apo-siderophore from the receptor. The loss of either TonB or the proton motive force in the transferrin receptor causes irreversible binding of the transferrin, and greatly impairs dissociation of haemoglobin from the HpuAB receptor (Rohde and Dyer, 2004).

TonB is also believed to interact with periplasmic binding proteins, which are located at the periplasmic site of the beta-barrel and are involved in periplasmic passage to the inner membrane permease. The hypothesis is based on the observations that the *E. coli* siderophore binding protein FhuD can form complexes with TonB (Carter *et al.*, 2006), as well as BtuF, the periplasmic binding protein for vitamin B12 (James *et al.*, 2009).

Haem transporters in *Neisseria* – HmbR and HpuAB

In *Neisseria* spp*.,* two haemoglobin receptors have been previously characterised in some detail. HmbR is a single component TonB-dependent transporter which is involved in haem acquisition from haemoglobin. Expression of this gene is associated with the invasive phenotype of *Neisseria meningitidis,* emphasising its role in iron acquisition in the blood stream (Harrison *et al.*, 2009). In gonorrhoeae, *hmbR* is present usually as a pseudogene, and no expression is observed (Stojiljkovic *et al.*, 1996). HpuAB is a bipartite receptor for both haemoglobin and haemoglobin-haptoglobin, and is expressed in pathogenic and commensal *Neisseria* spp*.* Despite low sequence homology between HmbR and HpuB (28% identity), and other characterised haem transporters, they share a typical structural topology for TonB dependent transporters. In addition, loop 7 in most haem/haemoglobin receptors contains

two highly conserved motifs, FRAP and an NPNL box, and a conserved histidine residue (Fig. 1.8). This region has been shown to be essential for haemoglobin binding and/or extraction of haem from the haemoglobin (Bracken *et al.*, 1999; Fusco *et al.*, 2013; Liu *et al.*, 2006). Deleting this region in the HmbR receptor did not affect haemoglobin binding to the receptor, but affected utilisation of haem from the haemoglobin, whereas growth supported by haem as a sole iron source was unaffected. This indicates that the region is important for haem extraction from its proteinacious complexes.

```
hmbR    AAQL-NQAWRVGYDITSGYRVPNASEVYF-TYNHGSGN---------WLPNPNLKAERST
hpuB    DWRF-TKHLHLLAKYSTGFRAPTSDETWL-LFPHPDFY---------LKANPNLKAEKAK
hasR    AVRPGVEWLELYTIYGKSWRPPAITETLTNGSAHSSST---------QYPNPFLQPERSR
hmuR    ---K-CSHVTNRLSYAEGYRAPSLQEMYF-FFNHGAFF---------IYGNPDLKPEKSR
hemR    SVTP-TDWLMLFGSYAQAFRAPTMGEMYN-DSKHFSMNIMGNTLTNYWVPNPNLKPETNE
```

Figure 1.8 Sequence alignment of fragments from five different TonB-dependent haem transporters.

Five haem TBDTs from various Gram-negative bacteria were analysed by Clustal Omega and Boxshade server: HAHmbR (Accession Number: NP_274673.1) and HpuB (AAC44893.2) *from Neisseria meningitidis*, HasR from *Serratia marcenscens* (CAE46936.1), HmuR from *Porphyromonas gingivalis* (YP_004509477.1) and HemR from *Yersinia enterocolitica* (CAA48250.1). The motifs FRAP and NPNL, and the conserved histidine residue are highlighted in red.

Unlike other haemoglobin receptors, HmbR also requires an intact plug domain for the utilisation of haem from haemoglobin. However, the histidine in position 500 which is highly conserved between haem binding receptors (Fig. 1.8) was not involved in haem transport in HmbR (Perkins-Balding *et al.*, 2003). In addition to loop 7, loop 6 was also shown to be involved in haem utilisation from haemoglobin, and showed the lowest sequence diversity when comparing numerous neisserial isolates (Evans *et al.*, 2010). Haemoglobin binding sites in HmbR were identified as loops 2 and 3, which show a high degree of sequence polymorphisms. This can explain the fact that HmbR, despite preference for human haemoglobin, can also bind haemoglobin from other sources (Evans *et al.*, 2010; Perkins-Balding *et al.*, 2003). Similarly to HmbR, HpuB also shows

high sequence conservation in loops 6 and 7 with high sequence divergence in the loop that is homologous to loop 2 (Harrison *et al.*, 2013). As mentioned before, the lipoprotein HpuA is essential for the utilisation of iron from haemoglobin and haemoglobin-haptoglobin complexes, and is probably responsible for the interactions with haptoglobin. This would explain the presence of an additional lipoprotein for the utilisation of haem from the more complex Hb-Hp substrate compared to HmbR, which is unable to utilise it (Rohde and Dyer, 2004).

Another TonB-dependent, single outer membrane zinc receptor, ZnuD (previously known as TdfJ), might also be involved in haem acquisition (Stork *et al.*, 2010). This receptor is able to support haem-dependent growth when expressed in the *E. coli* haem auxotroph strain. ZnuD is similar to other haem receptors in that it has the FRAP/NPNL motifs and the conserved histidine residues in the loop 7. The zinc binding site is located in loop 2. It is possible that this receptor might be involved in the transport of two different substrates depending on the circumstances and substrate availability. This is because its expression is regulated by two regulatory proteins: Zur and Fur (Kumar *et al.*, 2012).

Siderophore transporter – FetA

The ferric-xenosiderophore complexes have been shown to interact with FetA, a ferric enterobactin transporter (formerly known as FrpB) found in *N. gonorrhoeae.* This indicates that FetA may be involved in hijacking iron from exogenous iron complexes (Carson *et al.*, 1999). Other ligands for this receptor have been identified, and include salmochelin and dihydroxybenzoylserin (DHBS). This indicates broader ligand specificity (Hollander *et al.*, 2011). The *FetA* gene is located upstream to its periplasmic binding protein *fetB*, which has been shown to be indispensable for iron utilisation from siderophores. Downstream of the *fetB* gene is a designated ABC transporter FetCDEF (Carson *et al.*, 1999). *FetA* undergoes phase variation and is subject to iron fur-mediated regulation. The presence of a 91% identical FetA homolog in *N. meningitidis* suggests that these bacteria can also acquire iron in a similar way (Pettersson *et al.*, 1995).

1.2.6.4 Other uncharacterised iron transporters in *Neisseria*

Bioinformatic analysis of the genomes of *Neisseria* spp*.* resulted in the identification of several genes that display features typical for the TonB dependent transporter family (Tdfs) (Hagen and Cornelissen, 2006; Turner *et al.*, 2001). Further analysis of homology based on the consensus sequences revealed high homology of one of them, TdfF, to the siderophore receptor FpvA from *E. coli.* Two other genes,TdfG and TdfH, showed high homology to the haem TonB-dependent transporter HasR from *Serratia marcescens* (Turner *et al.*, 2001).

TdfF has been shown to be essential for intracellular iron acquisition by *Neisseria* spp., because a mutant with a defective TdfF protein was unable to survive within epithelial cells. The addition of excess free iron restored the intracellular survival of the mutant (Hagen and Cornelissen, 2006). Another argument that supports TdfF involvement in iron acquisition is the presence of the putative periplamic binding protein gene, with high similarity to *fetB,* upstream of the *tdfF* gene. TdfF expression is believed to be regulated in an iron-independent manner, as no studies have reported protein expression in conditions with different iron availability (Cornelissen and Hollander, 2011), and expression in gonococci was only observed when grown in the presence of serum (Hagen and Cornelissen, 2006). The involvement of another regulatory protein, MpeR (which belongs to the AraC family of proteins), in the regulation of expression is also speculated (Cornelissen and Hollander, 2011). This gene is also associated with virulence in *Neisseria* spp*.,* as the presence of the operon associated with this gene is unique to the pathogenic species.

TdfG and TdfH encode large proteins (>100 kDa) with features typical for TonB dependent haem transporters. However, mutants that expressed inactivated TdfG or TdfH retained the ability to grow on haem as a sole iron source, leaving the question about their function unanswered (Turner *et al.*, 2001).

1.2.6.5 TonB independent transport across the outer membrane

Although most iron and iron complexes are transported across the outer membrane via TonB-dependent receptors, *N. gonorrhoeae* can utilise iron from some siderophores or haem in a TonB-independent manner. Mutants lacking

the expression of *tonB* and *fetA* genes were still able to grow in the presence of xenosiderophores. This suggests that other structures in the outer membrane might be involved in siderophore or haem transport across the outer membrane (Strange *et al.*, 2011). One hypothesis mentions a possible role for PilQ in the passive transport of haem. PilQ is a component of the TypeIV secretion system that is responsible for the secretion of pili subunits across the outer membrane. A study performed by Chen *et al.* demonstrated that PilQ allows the passive passage of free haem, and some other compounds, in the periplasm. A phenotype with a single point mutation results in the increased leakage of low molecular weight compounds through the pore formed by PilQ (Chen *et al.*, 2004). However, studies performed by Strange *at al.* showed that TonB-independent transport of some xenosiderophores is still observed in mutants that lack the PilQ and Mtr proteins (protein forming a pore in the efflux pump) (Strange *et al.*, 2011). This indicates that some outer membrane porins might allow some passive transport through the outer membrane, or that there is an additional mechanism of transport which needs to be characterised.

1.2.6.6 Periplasmic transport of iron

Following passage through TbpA, free ferric iron is captured by the FbpA protein, which is attached to the periplasmic face of TbpA. It has been shown that TbpA shows higher affinity for the apo-FbpA than for the holo-FbpA, allowing dissociation of the protein after binding of iron. This higher affinity for iron seen in FbpA compared to TbpA is important for the unidirectional transport of iron from the outer membrane receptor into the periplasm (Siburt *et al.*, 2009).

FbpA is also known as a bacterial transferrin due to its structural and functional similarities to human transferrin. The ferric iron binding site in FbpA consists of four amino acids in a similar conformation as that seen in transferrin, and they both require the presence of an exogenous anion for tight iron binding. Additional similarities include the presence of a bilobial structure and similar mechanisms of ligand binding described as a Venus Fly Trap or Pac-Man motion, where two large lobes of the protein close around bound ligand. (Parker Siburt *et al.*, 2012).

FbpA is expressed from the fbpABC operon, which also encodes an inner membrane permease involved in the transport of iron in the cytoplasm (FbpB), and an ATP-binding protein that provides energy for translocation through the cytoplasmic membrane (FbpC). This operon was demonstrated to be essential for the utilisation of iron from other iron sources, like lactoferrin and some bacterial xenosiderophores. Bacteria without an FbpABC transport system were unable to survive when iron, in the form of lactoferrin, transferrin or siderophores, was used as the only iron source, but the growth was unaffected in the presence of haem. This indicates that haem transport across the inner membrane utilises an additional transport system yet to be identified (Strange *et al.*, 2011).

1.2.6.7 Haem metabolism in the cytoplasm

The key component in cytoplasmic haem metabolism in *Neisseria* is haem oxygenase (HO). Mutants of *N. meningitidis* and *N. gonorrhoeae* with an inactive haem oxygenase gene were unable to grow when haem or haem complexes were the sole iron source (Zhu *et al.*, 2000a). Haem oxyganase proteins are 95-98% identical in *Neisseria* spp*.,* with the residues that are involved in catalytic function being totally conserved. It also shares approximately 20% sequence identity with the human haem oxygase 1 (HO-1), with the highest homology in the sites important for HO function. The *hemO* gene is located upstream of the *hmbR* gene that encodes a haemoglobin receptor. Although both genes have their own promoters, which are regulated by the presence of an individual fur region, transcriptional coupling of both genes has been observed (Zhao *et al.*, 2010).

Haem oxygenase requires NADPH and O_2 for its activity. The end products of the reaction are: free reduced iron (Fe^{2+}), biliverdin and CO. Free reduced iron in the cytoplasm is then incorporated into different enzymes, or stored in the cytoplasm in the form of bacterioferritin. It is not clear how *Neisseria* spp*.* utilises either biliverdin or CO (Zhu *et al.*, 2000b).

1.2.6.8 Regulation of the gene expression involved in iron acquisition

A study by Grifantini identified 233 genes in the neisserial genome that are differentially expressed in iron-replete and iron-deplete conditions. Almost half

of them were predicted *in silico* to contain a Fur box in the promoter region. In *Neisseria* spp*.,* the Fur box consists of the consensus region nATwATnATwATnATwATn. The binding of the Fur protein to the promoter region has been experimentally confirmed for most of these predicted genes (Grifantini *et al.*, 2003).

Fur protein is highly conserved in *Neisseria*, with a sequence identity of over 90%. Between different species of Gram-negative bacteria there is quite a high sequence diversity; however, they do share some common features. The N-terminus contains an alpha-helical rich region that is involved in DNA binding, whereas the C-terminus is responsible for metal binding through its conserved histidine rich region, and forms the dimerization interface (Fillat, 2014).

The mechanism of Fur-mediated repression of gene expression has been widely studied, and it is characteristic for most of the iron receptors in *Neisseria* spp*.* In iron-replete conditions, binding of iron to Fur monomers triggers a conformational change in the proteins, which allows the formation of a dimer. An iron-loaded dimer binds to the Fur box in the promoter region and prevents binding of RNA polymerase, thus repressing protein expression. In contrast, in low iron conditions, Fur is prevalent in its monomeric form, which does not bind to DNA, so protein expression is unrestricted (Fig. 1.9.A.) (Escolar *et al.*, 1997). Studies by Delany and coworkers identified 83 genes regulated by Fur, of which only 44 genes were repressed. Surprisingly, 39 genes were activated, indicating a more versatile mechanism of gene regulation by Fur (Delany *et al.*, 2006). In the last few years, many studies have demonstrated that Fur-modulated protein expression by iron-independent mechanisms also required the involvement of other regulatory proteins and sRNAs (Yu and Genco, 2012).

Some bacteria are able to adapt the expression of their iron receptors in response to the presence of host-iron binding proteins or siderophores via the extracytoplasmic sigma factor (ECF). The HasR haemoglobin receptor from *Serratia marcenscens* is involved in signal transduction to the ECF sigma factor upon binding of its haemophore HasA. This results in the expression of HasR.
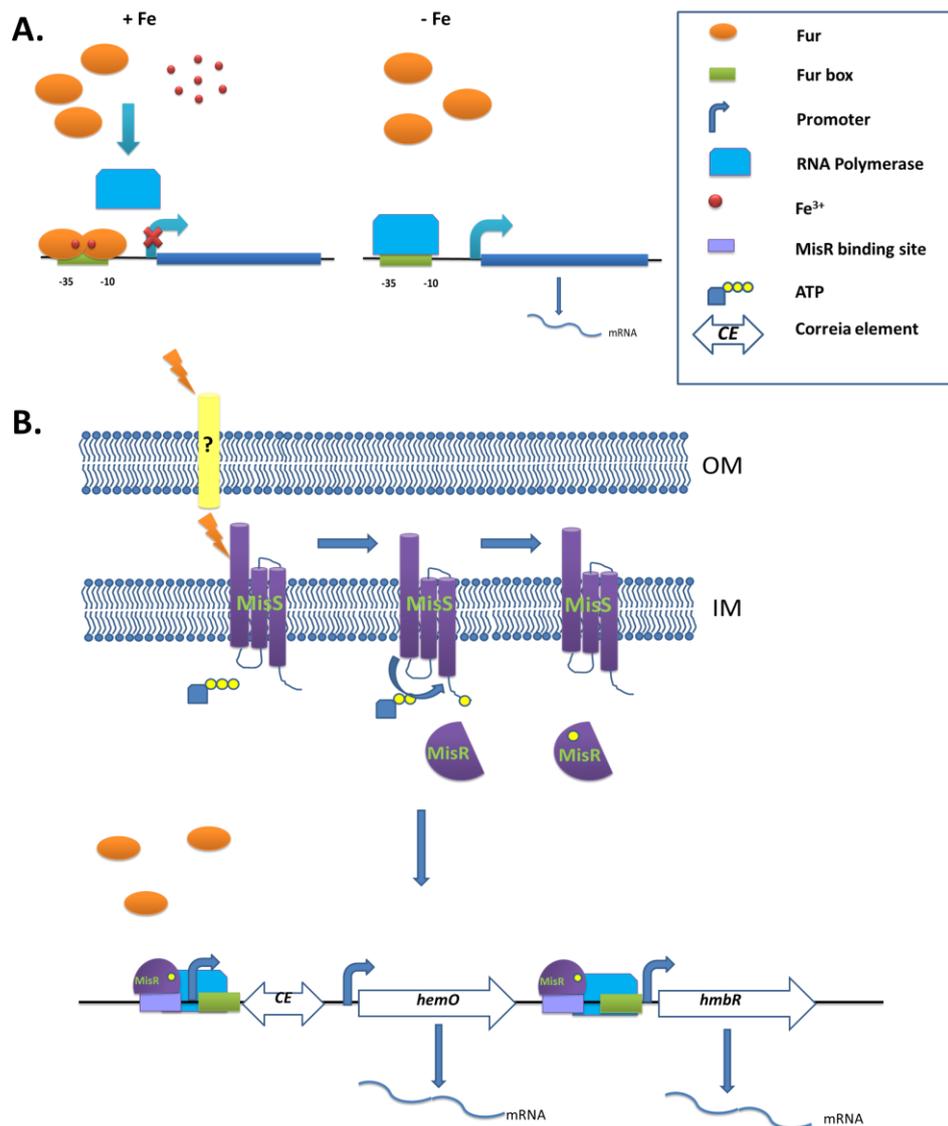
Figure 1.9 Regulation of gene expression in *Neisseria* spp. in response to the environmental factors.

A. Fur-mediated repression of genes involved in iron uptake and transport in iron-replete conditions. In the presence of iron, Fur forms a dimer which is able to bind to the Fur cassette in the promoter region, preventing attachment of the RNA Polymerase, and gene expression. In the absence of iron, Fur is present as a monomer which is unable to bind to FUR cassette, thus gene expression is unaffected (Figure adapted from Yu and and Genco with permission (Yu and Genco, 2012). **B.** The MisR/S two component system regulation of *hmbR* and *hemO* as a control mechanism is independent of Fur. MisS is a membrane-bound histidine kinase which undergoes auto-phospharylation in response to environmental signals. The phosphoryl group is then transferred to the cytoplasmic gene expression regulator MisR, which is a DNA binding protein. MisR has been shown to bind to the promoter regions of *hemO* and *hmbR* genes, thus increasing their expression levels. Figure based on (Tzeng *et al.*, 2006; Zhao *et al.*, 2010).

A study by Jordan and Saunders that compares the transcription profiles of *Nesseria meningitidis* when grown on different iron-binding host proteins (haemoglobin, transferrin and lactoferrin) revealed differences in the response and gene expression. These differences were not limited to the genes involved in the iron metabolism. Phenotypes grown on lactoferrin were associated with the up-regulation of genes involved in carriage, including genes involved in cell adhesion and oxidative stress, which are required for mucosal colonisation. Growth on transferrin resulted in the upregulation of the fewest number of genes, with the TbpA gene being up-regulated the most. The transcriptome of *N. meningitidis* grown on transferrin is similar to the expression profile of bacteria grown in iron-replete conditions. Responses to haemoglobin showed the up-regulation of genes that are typical of iron-replete conditions, and genes associated with an invasive phenotype and increased resistance to the host immune defence in the blood stream. The differential expression of the genes in response to different host iron-binding proteins is a mechanism of adaptation to different niches (Jordan and Saunders, 2009). It is not clear how bacteria are able to recognise changes in the extracellular environment and how the signals form the environment are translated into differential gene expression. The results of the study suggest that *Neisseria* may employ an unknown mechanism similar to the ECF sigma factor that is employed by iron host protein binding protein receptors.

Neisserial outer membrane receptors and other proteins involved in iron metabolism are regulated by iron-dependent fur-mediated mechanisms. Only TdfH expression is independent of the iron levels, and ZnuD is activated by iron (Cornelissen and Hollander, 2011).

An additional mechanism for the expression of the Hmbr receptor was described that involves the two-component MisR/S system. MisR/S is believed to be responsible for sensing of environmental stimuli and the transduction of the signal inside the cells. This allows the pathogen to adapt protein expression to the changing environment, for example adhesion to the host epithelial cell, intracellular survival (Jamet *et al.*, 2009; Tzeng *et al.*, 2006). The *hmbR* gene is located downstream of the *hemO* gene, and both genes have their own promoters and sequences typical for a Fur box, although they seem to be

transcriptionally linked. MisR/S is believed to activate expression, as mutants that do not express the MisR/S genes showed reduced levels of expressed *hmbR* and *hemO* genes, even in the iron-depleted conditions (Fig. 1.9.B.). Expression was then restored following complementation with MisR/S (Zhao *et al.*, 2010). MisR/S regulation seems to be independent of iron regulation, and might be an additional mechanism for controlling protein expression to allow adaptations to the changing environment. The *HpuAB* gene was also shown to be regulated by MisR/S, unlike the transferrin and lactoferrin receptors, which indicates that this regulation might be linked to haemoglobin utilisation.

1.2.6.9   Phase variation as a mechanism of regulating gene expression

In addition to regulating gene expression in response to environmental stimuli, *Neisseria* spp. modulates expression of many genes by stochastic changes in the gene sequence. This mechanism is believed to allow quick adaptation to the changing environment and immunogenic pressure before the microorganisms are subjected to selective stimuli. The genes subjected to phase variation are either in the ON or OFF state, where the protein is either expressed or not; or the level of protein expression is modulated by changes in the promoter region (Lovett, 2004). Genes undergoing phase variation are characterised by the presence of simple sequence repeats (SSR), which are composed of either poly(G) or poly(C), or with coding repeats (CR) which are fragments of four or more nucleotides. The SSR are prone to errors during DNA replication via slipped-strand mispairing. This results in frame-shift mutations, which in turn can prevent protein expression. The presence of SSR is also associated with translational frame-shifting (Alamro *et al.*, 2014) (Fig. 1.10).

Three of the outer membrane iron transporters (HmbR, HpuAB and FetA) were shown to be subjected to phase variation. HmbR and HpuAB genes contain poly(G) regions, and changes in the number of guanidine residues results in gene expression being in either an ON or OFF state (Lewis *et al.*, 1999). FetA phase variation occurs by targeting the poly(C)-rich region in the promoter region, resulting in changes in the protein expression between low and high levels (Fig. 1.10) (Carson *et al.*, 2000). TbpA does not undergo phase variation

as TbpA expression is necessary for neisserial survival within the host, and transferrin is the preferred iron source of the *Neisseria* spp*.* No evidence of phase variation of LbpAB was observed; however, the gene associated with the proteolytic cleavage of LpbA, NalP, undergoes phase variation. Thus, the function of LbpA can be modulated and contribute to immune system evasion via stochastic regulation of gene expression through NalP.
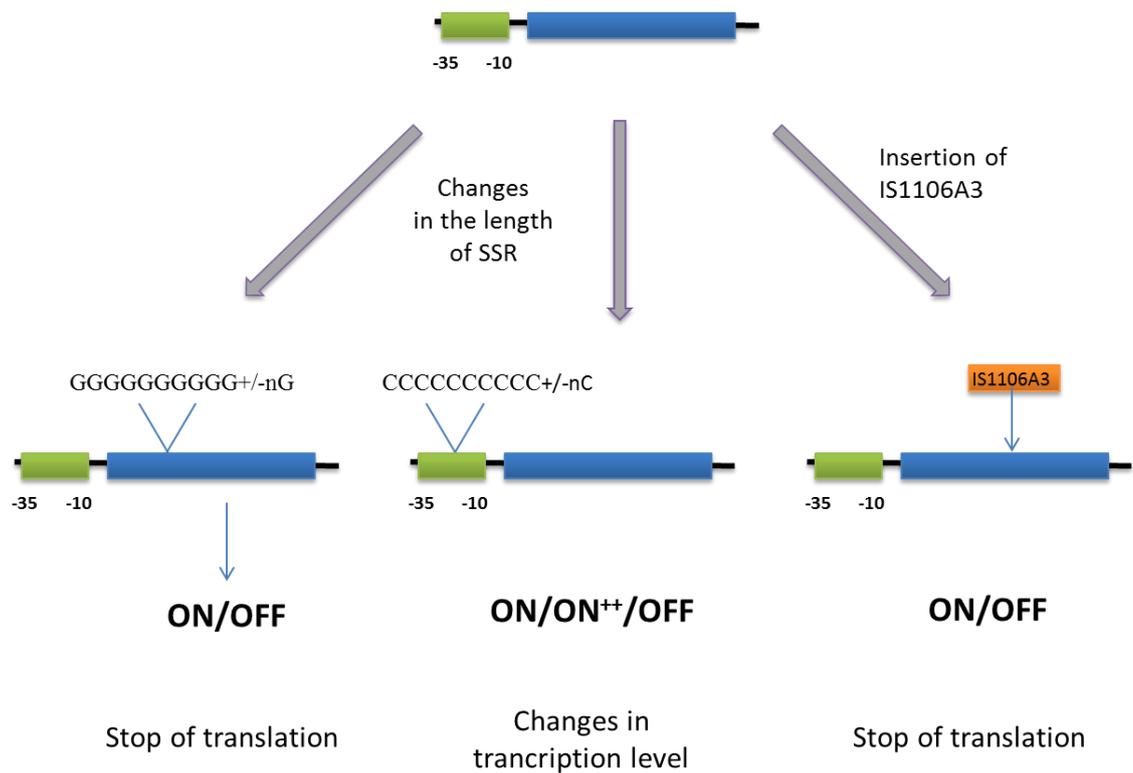
Figure 1.10 Schematic representation of the mechanisms of phase variation in *Neisseria*.

Changes in the length of simple sequence repeats (SSR) within the gene lead to ON/OFF switching of gene expression (for example *hmbR* and *hpuAB*, which contain poly-G tracts). However if these changes are in the promoter region, the length of the SSR determines levels of transcription (For example the *fetA* promoter, which contains poly-C tracts). Insertion or excision of mobile DNA fragments such as IS1106A3 can also result in ON/OFF switching by the introduction of pre-term Stop codons, as shown in *hpuA* (Tauseef *et al.*, 2011) Adapted from Lo *et al.* with permission (Lo *et al.*, 2009).

Phase variation of the outer membrane antigens favouring either OFF or low expression status has been reported in persistent carriage was strains, and is a result of the constant selective pressure from the host immune system (Alamro *et al.*, 2014). Differences in the prevalence of haemoglobin receptors HmbR and HpuAB that are in an ON state between carriage and invasive isolates has also been observed. The 91% of invasive isolates express either one or both haemoglobin receptors, with an overrepresentation of HmbR, in contrast to 71% of carriage isolates, which display preferential expression of HpuAB. Expression of either receptors or HmbR only seems to facilitate the invasion of *Neisseria* spp*.,* and can have an implication in immune evasion (Tauseef *et al.*, 2011). Phase variation is a form of adaptation to the different niches, with different availability of iron host proteins. This is because, during carriage, *N. meningitidis* resides in the nasopharynx where haemoglobin is scarce, whereas invasive isolates disseminate in the blood stream where levels of haemoglobin are higher, especially in later stages of infection.

In summary, *Neisseria* spp. have the ability to withstand host nutritional immunity when requires adaptation to different ecological niches during colonisation and invasion of the host. At each stage of neisserial infection the availability and source of iron can differ drastically. Therefore, *Neisseria* spp. has evolved a very sophisticated iron acquisition system which is tightly regulated during neisserial pathogenesis. An overview of all characterised up to date proteins involved in iron transport in *Neisseria* spp*.* is illustrated in Figure 1.11.
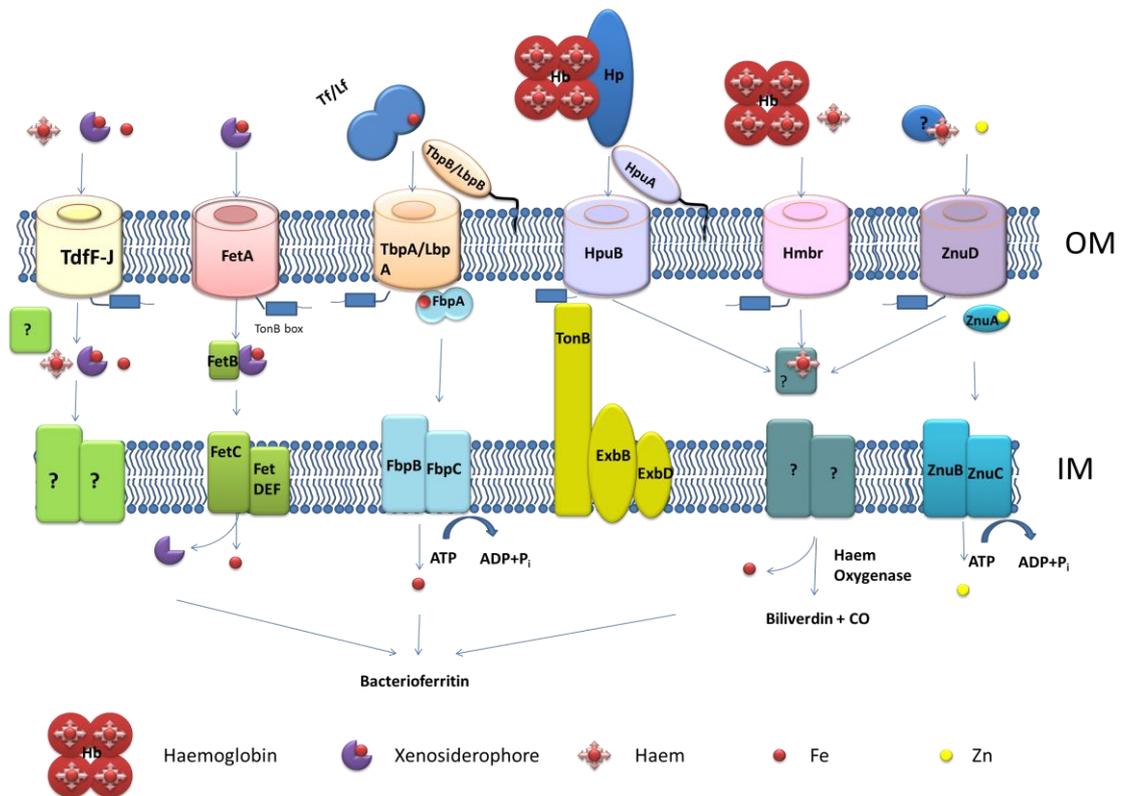
Figure 1.11 Overview of iron acquisition system in *Neisseria meningitidis*.

Iron in the host is sequestered by the host binding proteins transferrin (Tf) and lactoferrin (Lf), or it is bound to xenosiderophores produced by co-colonising bacteria. Most of the iron is available as haem complexes with host proteins, specifically haemoglobin (Hb), haptoglobin (Hp) and haemopoexin. Iron, in all forms, is transported through the outer membrane *via* TonB-dependent transporters (TBDT). The TBDTs either directly bind the ligand and transport iron or haem in the periplasm (single component TBDT) or require additional lipoprotein or haemophore for binding of the ligand (bipartite TBDT). To the single-component TBDTs belong: HmbR (haemoglobin receptor), FetA (siderophore receptor) Tdf transporters (transport of iron, siderophores and haem), ZnuD –zinc and haem transporter. The bipartite TBDTs are– HpuAB (haemoglobin-haptoglobin receptor), TbpAB – transferrin binding protein, LbpAB – lactoferrin binding protein or HpuAB – haemoglobin-haptoglobin binding protein. Periplasmic binding proteins (for example FbpA, FetA or an uncharacterised receptor) transport Fe, siderophores or haem respectively to the cytoplasmic membrane. Inner membrane translocators, specifically permease, which is responsible for both the removal iron from haem or siderophores and its translocation into the cytoplasm, and ATPase, which provides energy for transport. Iron in the cytoplasm is incorporated into the bacterial proteins, or it is stored as bacterioferritin. Haem, after iron removal, is metabolised by haem oxygenase to billiverdin and CO. Figure adapted from (Perkins-Balding *et al.*, 2004) (Kumar *et al.*, 2012) and Sathyamurthy (2011).

1.3 Introduction to study

1.3.1.1 Gly1ORF1 – a potential virulence factor of *Neisseria* spp*.*

A gene encoding Gly1ORF1 protein was discovered as a result of a search for potential cytotoxins produced by *Neisseria gonorrhoeae* (Arvidson *et al.*, 1999). In the study by Arvidson *et al.,* an *E. coli* library containing chromosomal fragments of *N. gonorrhoeae* strain MS11A was screened for clones with haemolytic activity, a property that often correlates with toxicity to other cell types (Bernheimer, 1988). The study resulted in the selection of one haemolytic clone with an insert referred to as gonolysin 1 (Gly1). The insert contained two open reading frames (ORF1 and ORF2) which were suggested to be co-transcribed probably generating a bicistronic transcript as the promoter sequence was found only upstream of the *gly1ORF1* gene. A ribosomal binding site was present upstream of each open reading frame (Fig. 1.12) (Arvidson *et al.*, 1999).
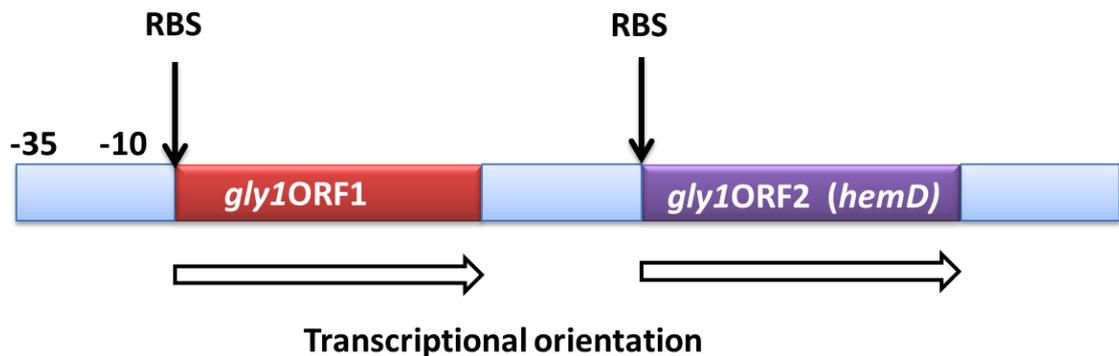
Figure 1.12 Schematic representation of the *gly1* locus.

The *gly1* locus consists of two open reading frames *gly*1ORF1 and *gly*1ORF2 separated by a short non-coding sequence. No promoter sequence has been detected between the ORFs suggesting the two genes are co-transcribed from the promoter upstream the ORF1.

Gly1ORF1 encodes a 140 aa protein with a predicted molecular weight of 17.8 kDa and pI of 9.57. The identified N-terminal signal peptide comprised of 21 amino acids is processed by the neisserial and *E. coli* signal peptide membrane proteases, resulting in the release of mature Gly1ORF1 product with the molecular weight of 15.6 kDa. Gly1ORF1 was predicted to have periplasmic

localisation with no obvious motifs indicating its insertion in the outer membrane. This was experimentally confirmed, with the data showing that Gly1ORF1 is present in the outer membrane of *N. gonorrhoeae* and *E. coli* constructs as well as in the filtered culture supernatant (Arvidson *et al.*, 1999). Initial structural studies indicate that Gly1ORF1 forms dimers in the solution (Arvidson *et al.*, 2003).
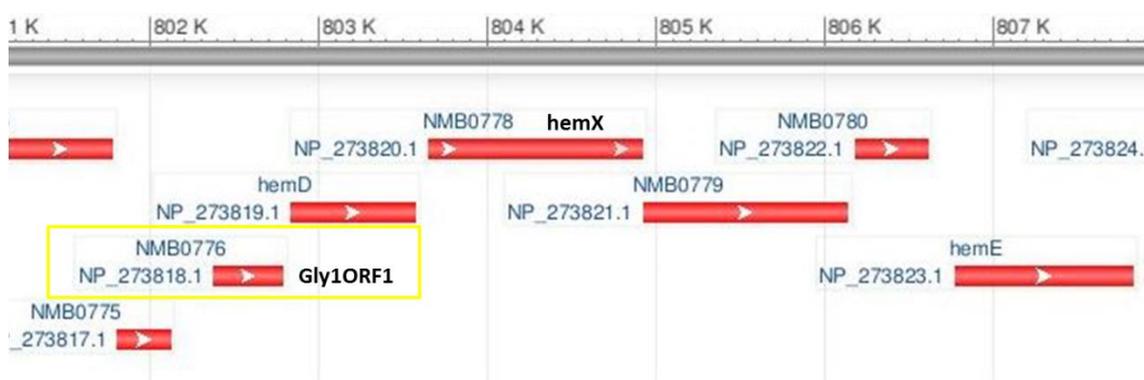
Figure 1.13 Genomic localisation of the *gly1ORF1* gene.

A snapshot of the region of the *N. meningitidis* MC58 genome (NCBI Reference Sequence: NC_003112.2) containing *gly1ORF1* gene (NMB0776), which is in close proximity to other genes predicted to be involved in haem biosynthesis. G*ly1ORF2 (hemD)* is located directly downstream gly1ORF1 and encodes a protein with high sequence homology to *E. coli* uroporphyrinogen III synthase, which is a key enzyme in haem biosynthesis. Gly1ORF1. The *hemX* gene (NMB0778) encodes uroporphyrin-III C-methyltransferase, whereas *hemE* encodes uroporphyrinogen decarboxylase.

Gly1ORF2 encodes a slightly acidic protein which is 247 aa long with a predicted molecular weight of 31.8 kDa and a pI of 6.25. The region between amino acids 86-187 has homology to the *hemD* gene of *E. coli* and *Psuedomonas aeruginosa* (27% and 36% identity respectively) encoding the enzyme uroporhirynogen III synthase that catalyses synthesis of uroporphyrynogen III from hydroxymethylbilane. Uroporphirynogen III is a precursor of the components of haem groups that are essential for synthesis of bacterial cytochromes and catalases (Granick and Beale, 1978). The genes *gly1ORF1* and *ORF2* were suggested to be co-transcribed (Arvidson *et al.*,

1999) (Fig. 1.13). This prompted speculations that Gly1ORF1 is also involved in haem metabolism.

Analysis of the distribution of *gly1* in *Neisseriae* by High Stringency Southern Blotting showed that the locus was present in three clinically most important species: *Neisseria meningitidis*, *N. gonorrhoeae* and N. *lactamica (Arvidson et al., 1999).* Meadows analysed the distribution of *gly1ORF1* within a number of carriage and invasive strains of *N. meningitidis* and *N. lactamica* and demonstrated that the gene is widespread in those species irrespectively of their invasiveness. The *gly1* gene is also conserved in *N. meningitidis* with over 90% identity (Meadows, 2004)

The studies on a *gly1ORF1* null mutant of *N. gonorrhoeae* showed that the product of this gene is not essential for the survival of *N. gonorrhoeae* in laboratory conditions. The absence of the *Gly1ORF*1 gene had no effect on the adherence to and invasion of three epithelial cell lines by *N. gonorrhoeae,* and only slightly decreased intracellular survival. However, the *gly1* null mutant showed increased toxicity to human fallopian tube tissue causing much more damage than the wild type strain. Complementation with recombinant Gly1ORF1 resulted in a phenotype similar to that of the wild type (Arvidson *et al.*, 1999). This prompted a suggestion that Gly1ORF1 could be responsible for modulation of cytotoxicity in *N. gonorrhoeae.*

Despite the fact that *Neisseria* spp*.* are normally non-heamolytic *in vivo*, the overexpression of *gly1ORF1* product resulted in the lysis of red blood cells on sheep or horse blood agar (Arvidson *et al.*, 1999). The haemolysis could be a result of the disturbance of the membrane integrity and subsequent leakage of some host proteins, i.e. lipases and proteases. The interaction of recombinant Gly1ORF1 with erythrocytes was also demonstrated by FACS analysis of RBC's incubated with fluorescently labelled recombinant Gly1ORF1 (B. Chen, M.Sc. Thesis, University of Sheffield) and by light microscopy showing that red blood cells incubated with Gly1ORF1 formed aggregates and drastically changed their typical morphology from a discoid to an irregular shape with lots of protrusion and blebs (Sathyamurthy, 2011).

Furthermore, the screening of a human cDNA library with Gly1ORF1 in a yeast two-hybrid system revealed that Gly1ORF1 interacts with erythrocyte membrane protein 4.2 (*EBP42)*, a major component of the RBC cytoskeleton (Meadows, 2004)). The interaction of Gly1ORF1 with EPB4.2 was further confirmed by Sathyamurthy by fluorescent microscopy. He demonstrated that fluorescently labelled Gly1ORF1 was binding to mammalian cells expressing recombinant EPB4.2 as opposed to cells containing an empty expression vector (Sathyamurthy, 2011).

The ability of Gly1ORF1 to interact with RBCs, together with the genetic localisation with genes involved in haem metabolism, pointed towards a possible role of Gly1ORF1 in haem acquisition. Following this lead, Sathyamurthy investigated the ability of Gly1ORF1 to bind haem. The results confirmed that Gly1ORF1 was indeed selectively sequestered from the solution and was binding to haemin-agarose beads. Also, when incubated with free haemin in solution Gly1ORF1 caused a spectrophotometric shift in the maximum absorbance wavelength of haem (Sathyamurthy, 2011).

In order to study Gly1ORF1's role in iron acquisition *in vivo*, a Gly1ORF1 isogenic mutant, *N. meningitidis* MC58, was generated. The growth rates of the wild-type (WT) and Gly1ORF1-*null-* mutant MC58 were measured in chemically defined minimal liquid media containing $FeSO_4$, haemin and haemoglobin as sole iron sources. The results showed that the Gly1ORF1 knock-out mutant is unable to grow in media with haem or haemoglobin as the sole iron source, whereas no difference in the growth rate of the WT and Gly1ORF1 knock-out was observed in the presence of $FeSO_4$. The growth of the ΔGly1ORF1 MC58 in haem- and haemoglobin-containing media was partially restored after complementation with recombinant Gly1ORF1. This indicates that Gly1ORF1 is involved in the acquisition of iron from haem or haemoglobin. The findings were further confirmed using a disc-diffusion growth assay, which also demonstrated limited growth of the Gly1ORF1 knock-out with haem and haemoglobin as the only iron source (Sathyamurthy, 2011).

Although the role of Gly1ORF1 in haem acquisition seems to be undisputable, the significance of this protein in neisserial pathogenesis is unclear. As

previously mentioned, the protein does not seem to be necessary for neisserial invasion and adherence ((Arvidson *et al.*, 1999). Also, it does not affect the intracellular survival of meningococci (Sathyamurthy, 2011). Unlike other proteins involved in iron acquisition, the expression of the protein is not regulated by iron levels in the Fur-dependent manner (Grifantini *et al.*, 2003). Gly1ORF1 upregulation is mentioned only in the study by Deghmane and colleagues which investigated the changes of protein expression upon contact with epithelial cells (Deghmane *et al.*, 2003).

However, the fact that Gly1ORF1 is involved in iron acquisition and that the *gly1*ORF1 gene is present in all tested invasive isolates and the majority of commensal isolates (Meadows, 2004) indicates that it might be an important virulence factor and needs further characterisation.

1.4 Project aims and hypothesis

Following hypotheses were proposed and explored in this study:

1. Gly1ORF1 is implicated in the haem acquisition in *Neisseria* through the interaction with host proteins and neisserial iron transporters.
2. Gly1ORF1 is able to elicit protective bactericidal response.
3. The newly identified Gly1ORF1 homologs protein family are a novel group of bacterial virulence factors with similar structure and functions.

The aims of the project:

1. To solve the crystal structure of Gly1ORF1 and identify the haem binding site.
2. To study Gly1ORF1 interaction with haem.
3. To further investigate Gly1ORF1 interaction with the red blood cell protein EPB4.2.
4. To investigate the interaction of Gly1ORF1 with neisserial proteins involved in iron acquisition.
5. To explore the Gly1ORF1 immunogenic properties and ability to elicit protective bactericidal response.
6. To characterise Gly1ORF1 homologs' and paralogs' structure and function.

CHAPTER 2   CHARACTERISATION AND CRYSTALLISATION OF GLY1ORF1 VARIANTS

2.1 Introduction

Although the *gly1ORF1* gene is highly conserved in *Neisseria* spp. and present in most of sequenced clinical isolates, the role of the protein in neisserial pathogenesis is not clear. Recent advances in the characterisation of the protein pointed towards the role of the protein in iron acquisition. As described in the introduction to the study, Gly1ORF1 has been shown to interact with red blood cells specifically targeting its major cytoskeleton protein EPB4.2 and bind haem.

The crystal structure of recombinant Gly1ORF1 of *N. gonorrhoeae* strain MS11 has been solved by Dennis Arvidson (Michigan State University, unpublished). The quaternary structure revealed the presence of a groove at the dimer interface which was suggested as a haem binding site according to molecular docking analysis performed by Prof. Jon Sayers. Although the groove seems to be a potential haem-binding site, several mutations introduced in this region by Sathyamurthy (Sathyamurthy, 2011) did not alter haemin binding *in vitro* as monitored spectrophotometrically. Also, no difference in the stoichiometry of binding was observed, as all mutated proteins demonstrated a 2:1 protein to haemin ratio. The results might indicate that the groove is not a binding site or the mutation did not cause sufficient steric hindrance to prevent haem binding.

This chapter describes studies on several new mutant versions of Gly1ORF1 of *N. meningitidis* MC58 that includes investigating binding of the haem and ability to oligomerise, as well as measuring haem binding affinity. The main aim of this chapter was to identify the haem binding site of Gly1ORF1 and includes attempts to crystallise Gly1ORF1 in the presence and absence of haemin.

2.2 Results and Discussion

2.2.1 Site Directed mutagenesis using the phosphorothioate approach

The wild type *gly1ORF1* gene, cloned into M13mp18 phage DNA, was subjected to site directed mutagenesis. This resulted in the creation of 7 single amino acid mutant genes, as described in the Materials and Methods Section 7.4.1. The aim of the mutagenesis described in this chapter was to identify the haem binding site of Gly1ORF1. Two possible binding sites were investigated: a groove at the dimer interface and the site between two β-sheets of a Gly1ORF1 monomer. Mutations of Glycine-47 to Arginine and Arginine-48 to Glycine were predicted to cause electrostatic and steric changes preventing formation of a dimer and no haem binding should be observed if the interface is a putative binding site (Fig. 2.1). The assumption that the haem binding site might be located between two β-sheets of a Gly1ORF1 monomer was based on the multiple sequence alignment of Gly1ORF1 homologs, which is described in Chapter 6 (Fig. 6.4 and Fig 6.5). The majority of the conserved residues (Phenylalanine-4, Lysine-12, Tyrosine-26, Phenylalanine-28, Glutamic Acid-35, Aspargine-67, and Tyrosine-72) are located in this position. Therefore, those residues were subjected to mutagenesis. Methionine-63 was selected due to its hydrophobic character and localisation (Fig. 2.2).
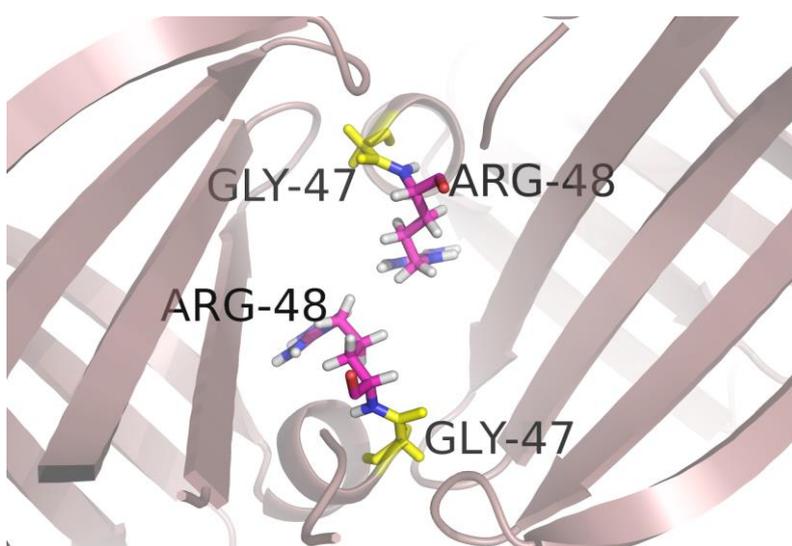
Figure 2.1 Gly1ORF1 residues at the predicted dimer's interface subjected to site directed mutagenesis.

Cartoon representation of the unpublished Gly1ORF1 structure of *N. gonorrhoeae* strain MS11 with the residues subjected to mutagenesis (Arginine-48 in magenta and Glycine-47 in yellow) shown as sticks. The image is rendered using Pymol.
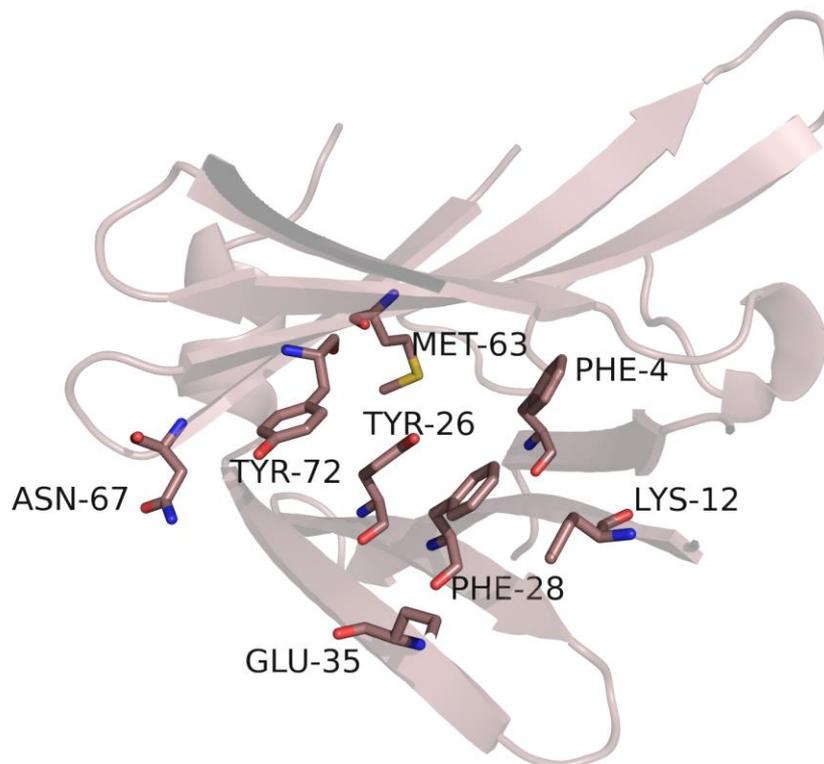
Figure 2.2 Residues subjected to site directed mutagenesis localised between two β-sheets of the Gly1ORF1 monomer.

Cartoon representation of the unpublished Gly1ORF1 structure of *N. gonorrhoeae* strain MS11 with the residues subjected to mutagenesis shown as sticks. Mutations were introduced using the phosphoraothioate and PCR based mothods described in the text. The image was rendered using Pymol.

Single-stranded circular phage DNA containing the *gly1ORF1* gene was annealed with oligonucleotides designed to introduce the required mutation. A polymerisation reaction with the addition of dCTPαS resulted in the formation of three circular forms of DNA: double stranded DNA (RFIV), with one wild type strand and one mutated strand, nicked double stranded DNA (RF II), and single-stranded DNA (Fig 2.4, Lane 1). Treatment with the T5 endonuclease targeted only the nicked and single stranded DNA, leaving the completely polymerised dsDNA intact (Fig. 2.4, Lane 2). Digestion using the restriction enzyme AvaI created a nick in a wild-type strand only as the mutated strand is resistant to AvaI activity due to the presence of phosphorothioate in the

cleavage position (Sayers *et al.*, 1992) (Fig. 2.4, Lane 3). The nicked wild type strand was removed by treating it with the T7 gene 6 exonuclease (Fig. 2.4, Lane 4). This left the mutated strand intact. After polymerisation with normal dNTPs, double-stranded DNA with the appropriate mutations on both strands was obtained (Fig. 2.4, Lane 5), and the mutated *gly1ORF1* gene was sub-cloned into the pJONEX4 C-histidine tagged expression vector, using the EcoRI and BamHI restriction enzymes. Sequencing results of the double-stranded M13mp18 phage DNA were obtained for all mutagenesis reactions, and confirmed the successful mutagenesis for all the mutants. Fragments were successfully sub-cloned into the pJONEX4 C-histidine tagged vector, and transformed into $CaCl_2$-competent M72 *E. coli* cells. The schematic representation of the site directed mutagenesis steps is shown in Figure 2.3.
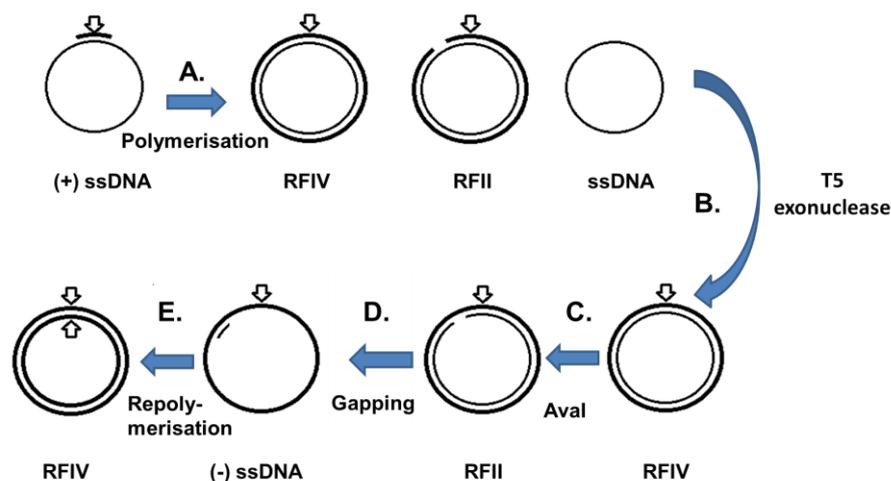
Figure 2.3 Schematic representation of the site directed mutagenesis using the phosphorothioate approach.

RFIV – circular double-stranded DNA; RFII – nicked circular double-stranded DNA; ssDNA – single stranded DNA. Arrows indicate an introduced mutation. Figure adapted from Sayers *et. al.,* 1992 (Sayers *et al.*, 1992), with permission.
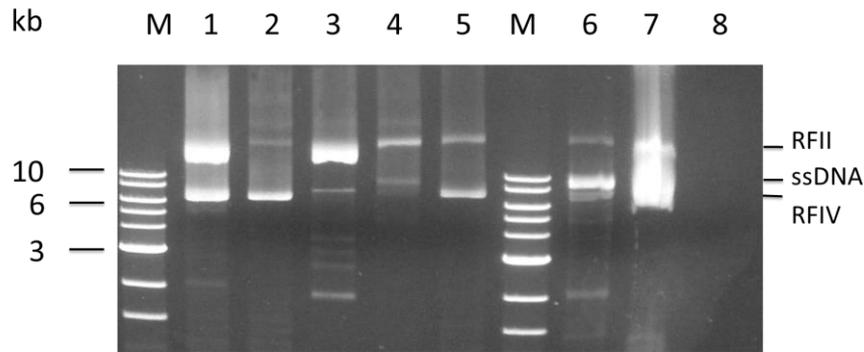
Figure 2.4 Successful mutation of R48 to G in Gly1ORF1 using phosphorothioate approach.

Site directed mutagenesis was performed according to the method illustrated in Fig. 2.3. Aliquots of the reaction at step A (**Lane 1**), step B (**Lane 2**), step C (**Lane 3**), step D (**Lane 4**) step E (**Lane 5**) and reference lanes with double-stranded (**Lane 6**) and single-stranded M13mp18Gly1 DNA (**Lane 7**) were applied to 1% agarose gel containing DTT next to DNA marker (**Lane M**) with sizes shown in kilobases. Negative control (**Lane 8**) contains a polymerisation sample without the addition of an oligonucleotide after treatment with T5 FEN endonuclease.

2.2.2 Small scale expression experiments of the genes encoding mutated Gly1ORF1.

Small scale heat induction of all Gly1ORF1 mutants was performed according to the protocol described in Materials and Methods Section 7.5.1. The heat induction was successful, with variable levels of relative expression (Fig. 2.5).
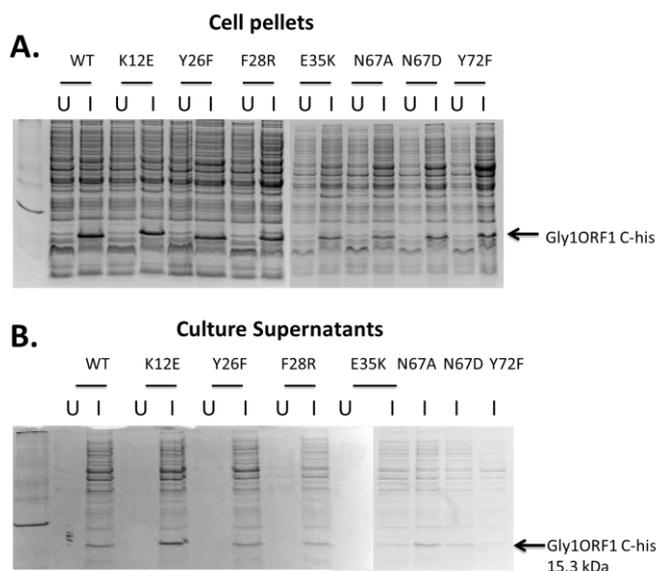
Figure 2.5 SDS-PAGE analysis of small scale expression of the Gly1 C-histidine tagged mutants by heat induction in M72 *E. coli* cells (Panel A) and supernatant (Panel B).

Samples before induction (**Lane U**) and after heat induction (**Lane I**) were separated on 13% SDS-PAGE gel and stained with Coomasie. Arrows indicate bands corresponding to Gly1ORF1 C-histidine tagged mutants.

74

Gly1ORF1 G47R and R48G expression was not visible on the SDS-PAGE gel (data not shown). However, western blot with anti-Gly1ORF1 rabbit polyclonal antibodies confirmed the presence of the protein in culture supernatants of tested clones (Fig. 2.6). The Gly1ORF1 protein in the cell pellet was mostly insoluble (data not shown). However, attempts to refold this could have resulted in purifying a protein that still contained a signal peptide. It was decided that only the secreted protein would be purified. Thus, a large-scale purification was performed using the culture supernatants from induced cells.
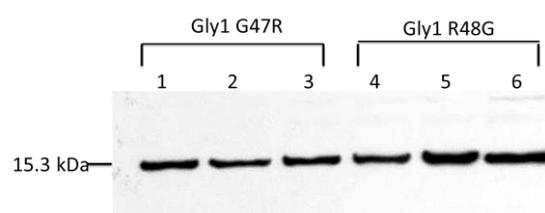
Figure 2.6 Western blot analysis of small scale heat induction.of Gly1ORF mutants G47R and R48G showing presence of the protein in the culture supernatants.

Three clones for each mutant were tested in parallel (**Lanes 1-3 and 4-6**)). Proteins were detected using rabbit anti-Gly1ORF1 polyclonal antibodies.

2.2.3 PCR based Site Directed Mutagenesis

Four more mutants were produced using a PCR based method (see Materials and Methods Section 7.4.2), where the whole plasmid (pJONEX4-Gly1ORF1 C-histidine tagged) was amplified using mutagenic primers. This resulted in multiple copies of the linear mutated DNA (Fig. 2.7, Fig. 2.8, Panel A). The linearised, mutated plasmid was treated with T4 polynucleotide kinase to phosphorylate the 5′ ends, and the template DNA was removed using DpnI, a restriction enzyme which cleaves only methylated DNA. Constructs were then circularised by ligation with T4 ligase (Fig. 2.6, Panel B.), and transformed into CaCl$_2$ - competent M72 *E.coli* cells. The schematic representation of the principle of this method is shown in Figure 2.7.
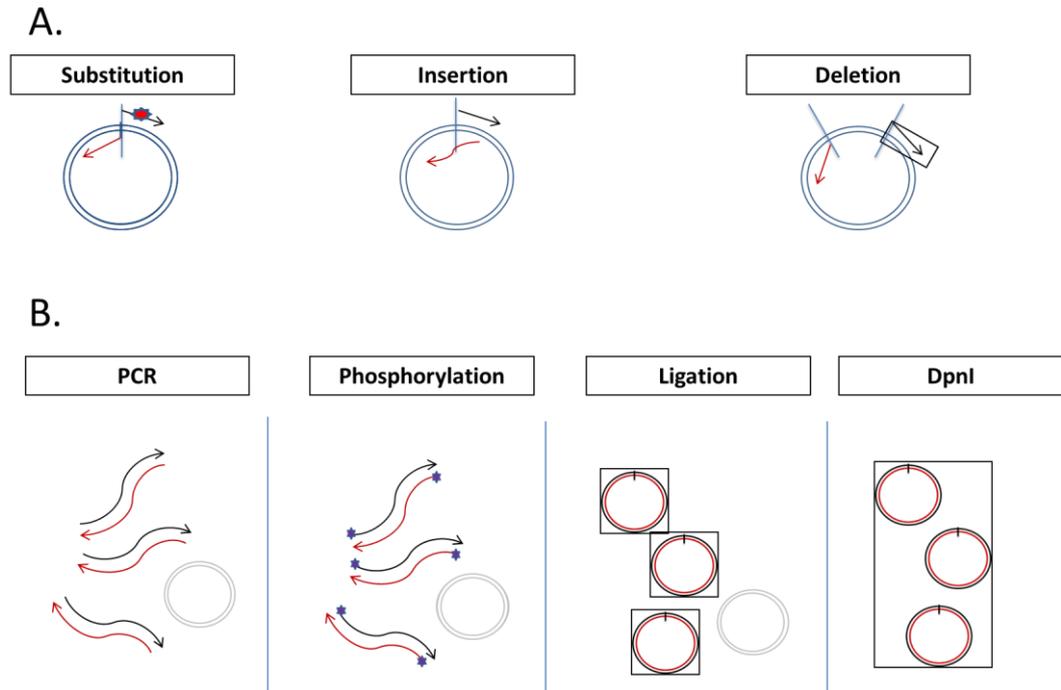
Figure 2.7 Schematic representation of the single step PCR based method.

A. Principle of primer design for different mutations. **B.** Steps of the mutagenesis. Grey circles illustrate template DNA plasmid. Black and red lines and circles are amplified mutagenized DNA in linear and circularised form respectively. See further explanation in the text (Picture adapted from the NEB website with permission).
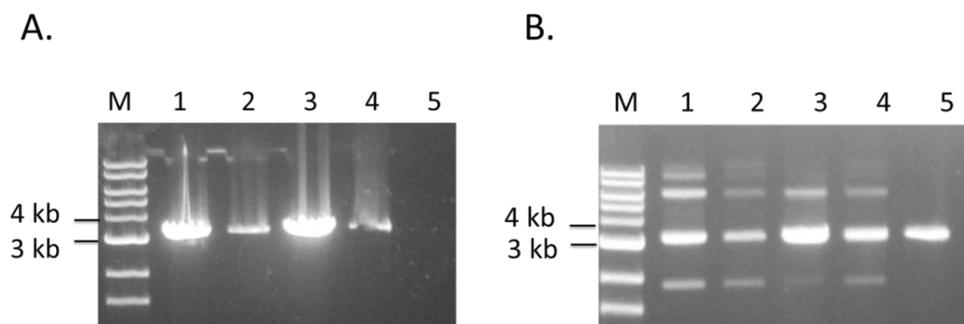
Figure 2.8 Agarose gel analysis of the site-directed mutagenesis by a single-step PCR method. showing PCR products of the linearised pJONEX4 C-histidine – Gly1ORF1 (Panel A) and ligated mutagenized plasmids (Panel B)

Panel A: Lane M – 1 kb DNA ladder, **Lane 1**- PCR product of the pJONEX4 C-histidine tag vector with the Gly1F4A mutant, **Lane 2** – Gly1K12A, **Lane 3** – Gly1 M63A, **Lane 4** – Gly1dSTOP, **Lane 5** – negative control of PCR reaction with no template and K12A primers. **Panel B:** Ligation of the linear mutagenized plasmids. **Lane 1- 4** – constructs in the same order as in panel A., **Lane 5** - negative control: construct Gly1 K12A without T4 ligase.

After successful transformation (Materials and Methods Section 7.3.5), DNA from three colonies for each construct were sequenced to confirm the presence of the appropriate mutation. All clones were tested for the ability to produce the mutated protein in 5 ml cultures (Fig. 2.9).
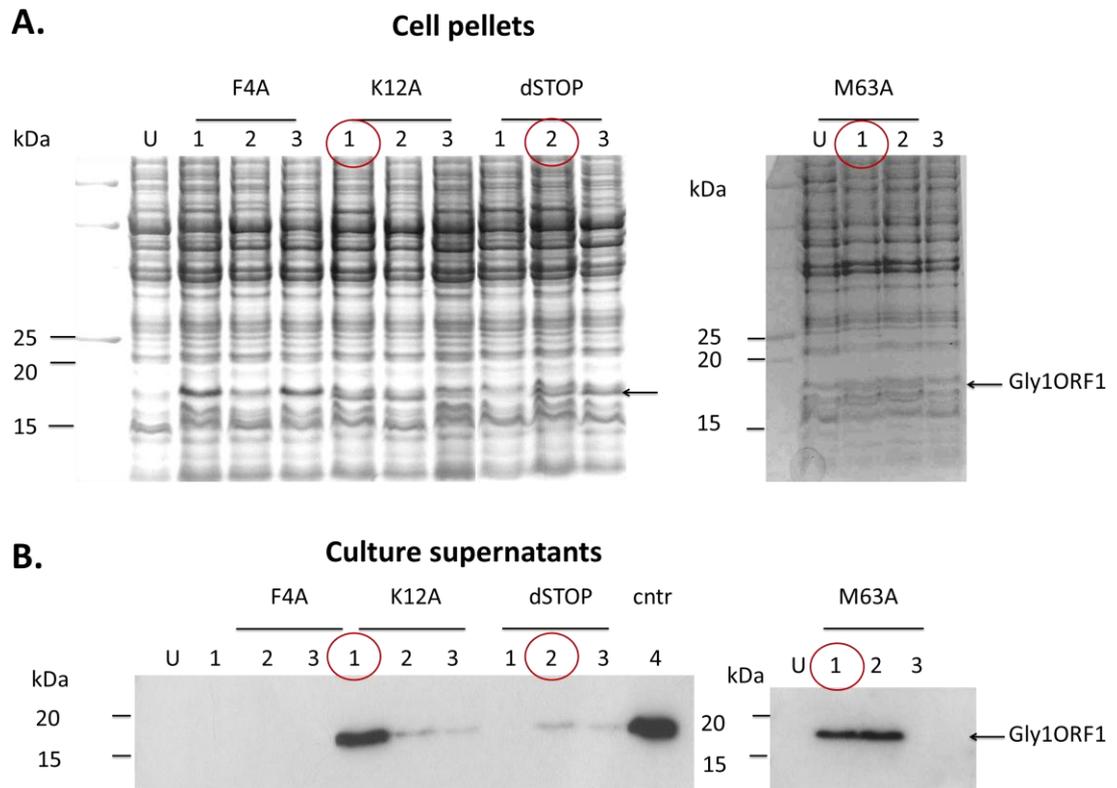
Figure 2.9 Detection of mutant Gly1ORF1 proteins production in cell pellets (Panel A) and culture supernatants (Panel B) using SDS-PAGE and western blot respectively.

Three clones (**Lanes 1-3**) of each mutant Gly1ORF1 (F4A, K12A, dSTOP, M63A) were heat induced in 5 ml cultures. The Gly1ORF1 mutant proteins in supernatants were detected using anti-Gly1ORF1 polyclonal antibodies. The un-induced Gly1ORF1 C-histidine tag cultures are shown in **Lane U**. Purified Gly1ORF1 C-histidine tag protein (**Lane 4**) was used as a positive control for western blot. Clones selected for large scale induction are marked with red circles.

There was no detectable protein production of the Gly1F4A protein in either cell pellets or culture supernatant in any of the three clones when tested by SDS-PAGE and western blot (Fig. 2.9). This occurred even though the sequencing results confirmed the correct sequence and presence of the mutation. It is probable that selected Gly1F4A clones contained mutation within the plasmid

DNA sequence that affected successful expression of the gene. Sequencing results confirmed the presence of the expected mutation in two of the three Gly1K12A clones. The production of the Gly1K12A-1 protein was higher than in the other K12A clones (Fig. 2.9, Panel B). Sequencing of the Gly1 dSTOP clones showed the correct sequence for two out of the three clones, but the clone 1 had a spontaneous mutation (resulting in a valine 19 being replaced by methionine), which resulted in a lack of protein production (Fig. 2.9). A frame-shift mutation was present in the Gly1M63A - clone 1, which explains the lack of expression of this construct (Fig 2.9, Panel B)

In summary, the single step PCR mutagenesis method is a quick and uncomplicated method to obtain a desired mutation. However, the KAPA Biosystems HiFi Hot Start polymerase, which was used in the PCR, has a low proofreading activity, which resulted in many spontaneous mutations. The mutations in the insert region were easily detected by sequencing, and incorrect clones were discarded. However, it is not common practice to sequence the whole cloning vector. Mutations in the vector sequence itself could affect expression of an insert or the antibiotic resistance selection genes. For large scale protein production, the following clones were selected: Gly1K12A 1, Gly1dSTOP 2 and Gly1M63A 1 (Fig. 2.9).

2.2.4 Large scale production of the WT Gly1ORF1 C-histidine tagged protein and the Ni-affinity purification of the protein

The large scale production of the proteins was carried out according to the protocol described in the Materials and Methods Section 7.5.2. Levels of produced Gly1 C-histidine tagged protein in the 3 litre fermenter differed from batch to batch. The induced supernatant, after overnight incubation, showed great variability in Gly1ORF1 concentration, and contamination with other proteins. In the picture shown below (Fig. 2.10) the Gly1 C-histidine tagged protein was readily detected in the supernatant; however the presence of many contaminating proteins indicated partial cell lysis.
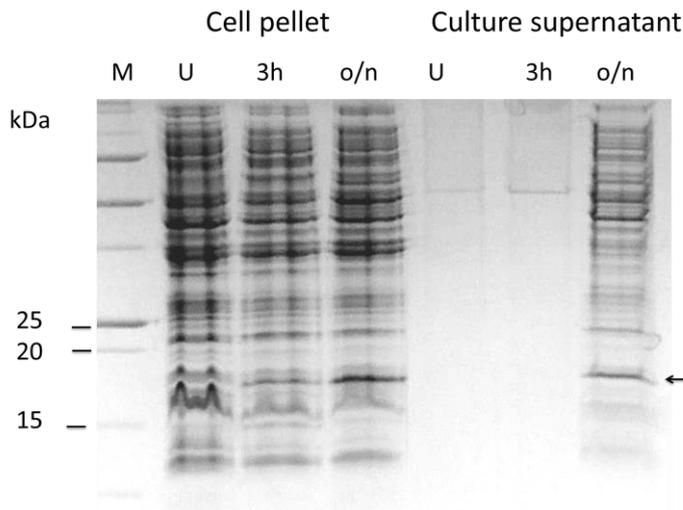
Figure 2.10 Large scale protein production of WT Gly1ORF1 C-histidine in 3 l culture with 4YT media.

Lane U – sample before induction; Lane 3h – sample 3 hours after induction; Lane o/n – sample after overnight incubation.

2.2.5 Purification of Gly1ORF1 on Ni-affinity column

Batch to batch variation was also observed when purifying the Gly1ORF1 C-histidine tagged protein on a Ni-affinity column (see Materials and Methods Section 7.6.1). Some purification attempts resulted in high yields of purified protein with little or no protein observed in the flow-through from the column. Other purification attempts, however, resulted in only partial binding of the protein to the Ni-affinity column, which resulted in very low yields of the protein. On other occasions the culture supernatant passed through the Ni-affinity column and removed the nickel ions from agarose beads, and no binding was observed.

It was hypothesised that the culture supernatants contained metal scavenging proteins or molecules which were removing nickel ions from the Ni-chelate resin, thus reducing the capacity of the column. The effect of addition of $NiCl_2$ to the culture supernatant on the yields of recovered protein was tested. Adding $NiCl_2$ before loading the culture supernatant on the column enhanced binding of the Gly1 C-histidine tagged protein to the matrix. After addition of $NiCl_2$ to the final concentration of 2.5 mM most of the Gly1ORF1 C-histidine tagged protein was bound to the column (Fig. 2.11, Panel A). Increasing $NiCl_2$ concentration up to 10 mM did not increase yields of recovered protein, as presumably at this

concentration it competes with the immobilized nickel on the column for the tagged protein (Fig. 2.11, Panel B).
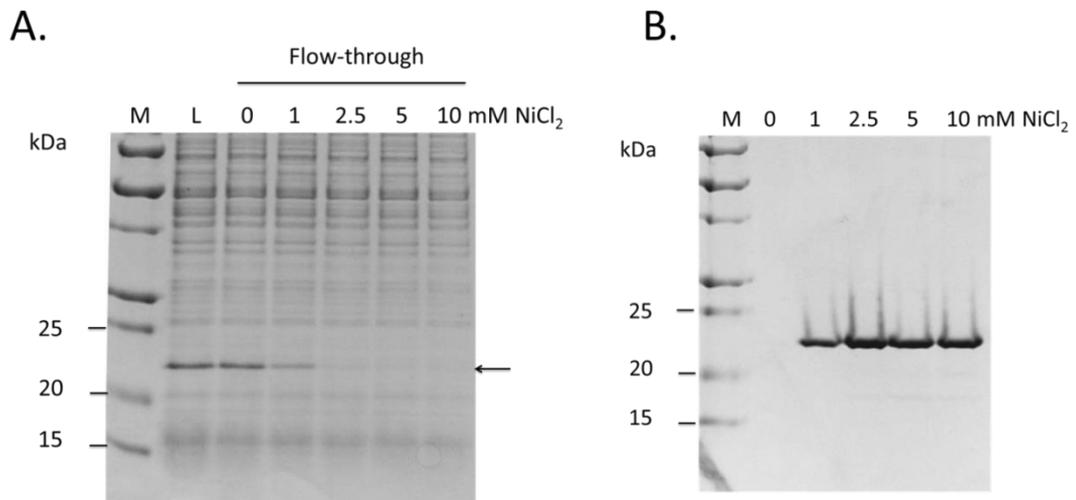
Figure 2.11 The addition of NiCl$_2$ to the culture supernatant aids the recovery rate of the Gly1 C-histidine tagged protein from the Ni--affinity column.

Supernatants containing secreted Gly1 C-histidine tagged protein (Panel. A., Lane L) were supplemented with increasing concentrations of NiCl$_2$ and loaded on the equilibrated Ni-affinity column. **A.** Flow-through samples analysed using SDS-PAGE, show almost complete binding of the tagged protein to the column at 2.5 mM NiCl$_2$. **B.** The Gly1 C-histidine tagged protein was eluted from the column using 500 mM imidazole, and the eluted fractions were analysed by SDS-PAGE.

Gly1ORF1, in the presence of NiCl$_2$, was purified on a larger scale using 1 litre of supernatant. Protein samples were dialysed into 25 mM NaOAc, pH 4; 50 mM NaCl. In low pH conditions Gly1ORF1 remains soluble even at the concentration up to 25 mg/ml (Fig. 2.12).
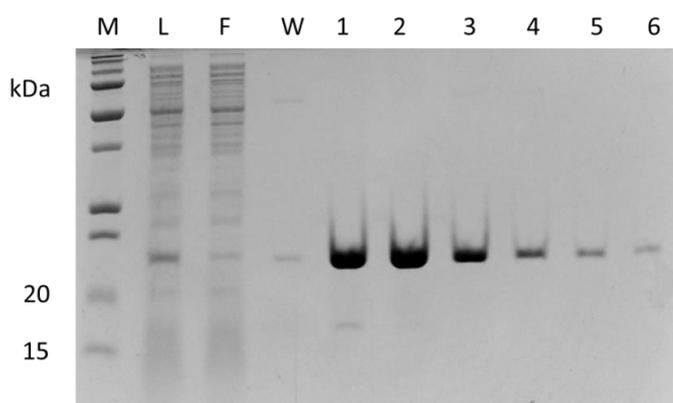
Figure 2.12 Purification of the C-histidine tagged Gly1ORF1 protein on a Ni--affinity column from the supernatant.

Lane L – sample before loading on to the column, **Lane – F**, Sample after passing through the column, **Lane W** – wash of the column with low imidazole concentration (25 mM), **Lanes 1 – 6** – eluted fractions.

Adding NiCl$_2$ to the supernatants was also effective for the purification of some of the mutant versions of the protein, specifically K12A and Y72F. These were purified from the supernatants at high yields (over 10 mg of protein per 1 litre of supernatant). However, the majority of the Gly1ORF1 mutant proteins did not bind to the Ni-resin despite the presence of NiCl$_2$. The proteins were only recovered from the supernatants after ammonium sulphate precipitation. One possible explanation is that the proteins were incorporated into outer membrane vesicles that were released by *E. coli* during protein induction. The histidine-tag would then be inaccessible, and unable to interact with the Ni-resin, which resulted in low binding to the Ni-affinity column.

2.2.6 Determining the size and distribution of the Gly1ORF1 C-histidine tagged multimeric forms using Dynamic Light Scattering

Dynamic light scattering methods enable the measurement of the size of molecules in a solution or suspension, as well their size distribution. In this method, the protein solution is exposed to light from a laser. The constant movement of the molecule causes the dispersion of light in different directions, which results in fluctuations in the absorbed light. The smaller the particles, the higher the fluctuations of absorbed light intensity are observed. From the intensity of these fluctuations, it is possible to calculate the velocity of the Brownian motion which, by using Stokes-Einstein relationship, enables the determination of the size of the molecule.

The WT Gly1ORF1 (0.5 mg/ml) was analysed using Dynamic Light Scattering (DLS) (see Materials and Methods Section 7.8.4). To disrupt the monomeric form, DTT was added to a final concentration of 10 mM, and the sample was measured by DLS to investigate changes in the oligomerisation state. The results identified three populations of peaks, with average sizes which are shown in Table 2.1.

Table 2.1 Results of the DLS analysis of Gly1ORF1 size in non-reducing and reducing conditions*

	Gly1ORF1	Gly1ORF1 + 10 mM DTT
Peak 1	6.82 nm ± 0.92	6.90 nm ± 0.81
Peak 2	30.95 nm ± 2.93	31.03 nm ± 2.58
Peak 3	180.83 nm ± 15.86	261.57 nm ± 26.42

*Results of three independent measurements with SEM using protein samples from the same preparation

There was no difference in the first two peaks, which correspond to the protein monomer and probably a dimer, between the non-reducing and reducing conditions. The largest oligomer showed a significant increase in size when in reducing conditions (Fig. 2.13).

These results indicate that the oligomeric forms of Gly1 are quite stable, and the presence of DTT did not disrupt the smaller oligomers. However, the largest oligomer expanded in size when in reducing conditions, which can be explained by the relaxation of the structure due to the loss of disulphide bonds.

A.

Size distribution Intensity

— 0.5 mg/ml Gly1ORF1

— 0.5 mg/ml Gly1ORF1 + 10 mM DTT

B.

Gly1ORF1 - analysis of peak sizes

n=3

p=0.002

p>0.999

p>0.999

■ 0.5 mg/ml Gly1ORF1

■ 0.5 mg/ml Gly1ORF1 + 10 mM DTT

C.

Gly1ORF1 - analysis of peak areas intensities

n=3

p=0.986

p=0.3321

p=0.1348

■ 0.5 mg/ml Gly1ORF1

■ 0.5 mg/ml Gly1ORF1 + 10 mM DTT
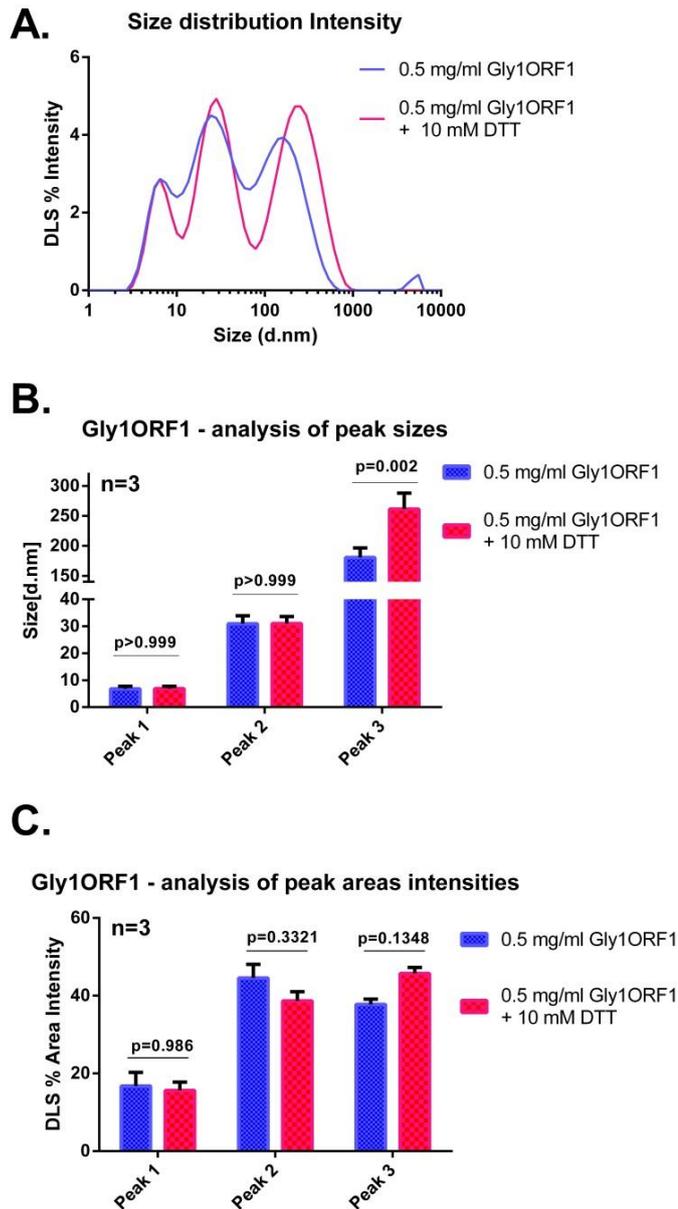
Figure 2.13 Analysis of the size and distribution of multimers of Gly1ORF1 by DLS.

A. Distribution of the size intensity of Gly1ORF1 in non-reducing (blue) and reducing conditions (pink). This shows the presence of three populations of proteins, with similar sizes for the two peaks of smaller size, and a slight increase in the size of the largest oligomer in reducing conditions.

B. Comparison of the sizes of the three populations of Gly1ORF1 multimers in non-reducing (blue) and reducing conditions (red) demonstrating a significant difference (p=0.002) in the size of the largest oligomer. The student t-test was used for statistical analysis.

C. Comparison of the percentage intensity of the peaks corresponding to each multimeric form. The percentage distribution of each multimeric form is similar between non-reducing and reducing conditions. In both cases, bars corresponding to the monomers are less abundant then the multimers, with similar distribution for the two larger oligomers.

With the monomers of Gly1ORF1 in a globular form, it would be expected that the dimers should be double the size of the monomer. However, in this experiment, a second peak that probably corresponds to a dimer is almost five times bigger than that of the monomer. Similarly, the third peak is also approximately five times the size of the second peak. Therefore, it was

suggested that Gly1ORF1 can form multimers at different interface than the one shown in Gly1ORF1 crystal structure. For example, it is possible that Gly1ORF1 can undergo a domain swap, leading to the formation of a chain of Gly1ORF1 monomers where the N-proximal β-sheet of one protein interacts with the C-proximal β-sheet of a second, and is stabilised by an intra-molecular disulphide bond. As a consequence, the size of the molecules measured by DLS would be much bigger as the method determines the diameter of the rotating molecule.

2.2.7 Haemin binding by the Gly1ORF1 mutant proteins

The Gly1ORF1 mutant proteins were produced and purified by Ni--affinity chromatography, similarly to the wild-type protein. In order to establish if the introduced mutations prevented haemin binding by the protein, a haemin-agarose pull-down was performed as described in the Materials and Methods Section (7.8.1).

All mutants examined retained the ability to bind to the immobilised haemin, with a visible enrichment of protein on the beads Fig. 2.14, Lane B), and a reduction in the amount of protein in the supernatant after incubation (Fig 2.14, Lane S). No such effect was observed for BSA, which was used in the assay as a negative control and to increase the solubility of Gly1ORF1 at neutral pH. Only faint bands corresponding to BSA were observed in the fractions containing proteins bound to the haemin beads (Fig. 2.14, Lane B). This is probably the result of non-specific interactions with the haemin. It is not clear whether the introduced mutations are located away from the actual haemin binding site, or if a single amino-acid mutation is not capable of causing sufficient steric hindrance to prevent the interaction. As this assay is only qualitative, no information is available about whether the mutated protein's affinity for haemin was altered.
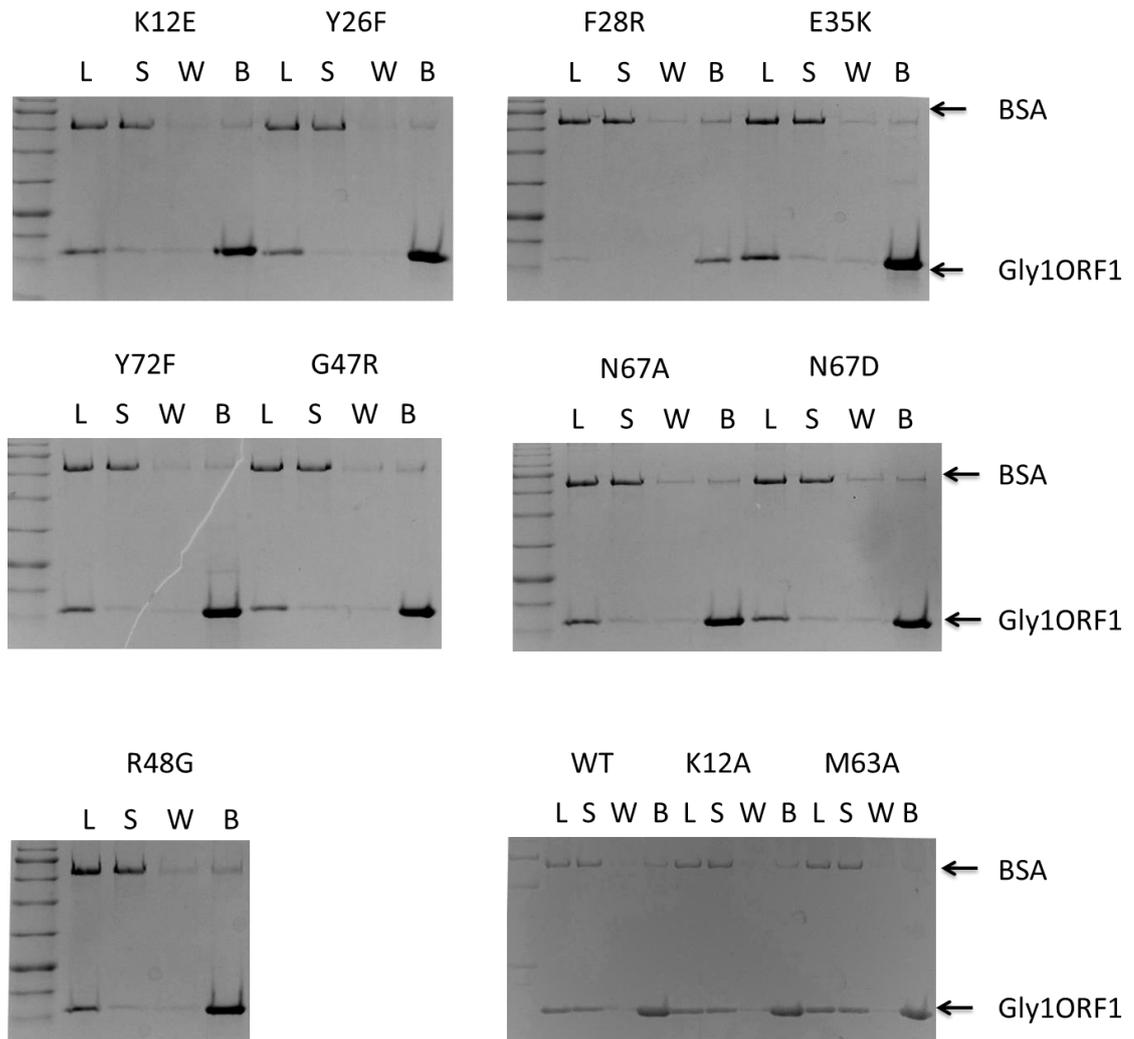
Figure 2.14 Selective binding of the mutated Gly1ORF1 proteins to haemin-agarose beads in the pull down assay.

Samples containing mutant and wild type (WT) Gly1ORF1 proteins and bovine serum albumin BSA (**Lanes L**) were incubated with haemin agarose beads. BSA served as a stabilising agent and a negative control for the assay. The haemin beads with bound proteins were centrifuged down and supernatant containing unbound proteins was removed (**Lanes S**) The beads were washed with PBS containing 0.5 M NaCl (**Lanes W**) to reduce nonspecific interactions with the beads. The proteins were eluted from the beads by boiling in SDS-PAGE loading buffer and added to SDS-PAGE gel (**Lanes B**).

2.2.8 Analysis of the oligomerisation of Gly1ORF1 by glutaraldehyde cross-linking

It was previously demonstrated that Gly1ORF1 is able to oligomerise by the formation of intramolecular disulphide bonds, as treating the protein with a reducing agent, such as DTT, disrupted the oligomeric forms back to monomers (Meadows, 2004). The data from the DLS experiment shown in section 2.2.6 also confirm this observation. Trapping of oligomers can be achieved by the addition of a cross-linking agent, such as glutaraldehyde. This amine-reactive homo-bifunctional cross-linker forms covalent cross-links between lysine side-chains that are in close proximity.

In order to examine if the introduced mutations caused any disruptions to the oligomerisation of the proteins, all purified mutants were treated with glutaraldehyde as described in Materials and Methods Section 7.8.5. The proteins were then separated using SDS-PAGE and visualized by western blot. The low solubility of the proteins at neutral pH, which is optimal for the reaction, precluded testing the proteins at higher concentrations. On the initial western blots, using standard transfer conditions, no proteins corresponding to the Gly1ORF1 oligomers were observed (data not shown). However, almost complete loss of the Gly1ORF1 monomers was observed for all tested proteins with increasing concentration of glutaraldehyde (data not shown). These results indicated that the Gly1ORF1 oligomeric forms were either too large to be separated by SDS-PAGE, or the transfer of the separated protein was ineffective. Reducing the concentration of the SDS-PAGE gel to 7.5% and increasing the transfer time allowed the detection of the multimeric forms for all the Gly1ORF1 mutants and the wild-type protein (Fig.2.15). All Gly1ORF1 mutants were able to form oligomeric forms. A quantitative comparison of the proteins was not possible due to the formation of aggregates that did not enter the gel, which is shown in the third blot, where the stacking gel was left intact (Fig. 2.15).
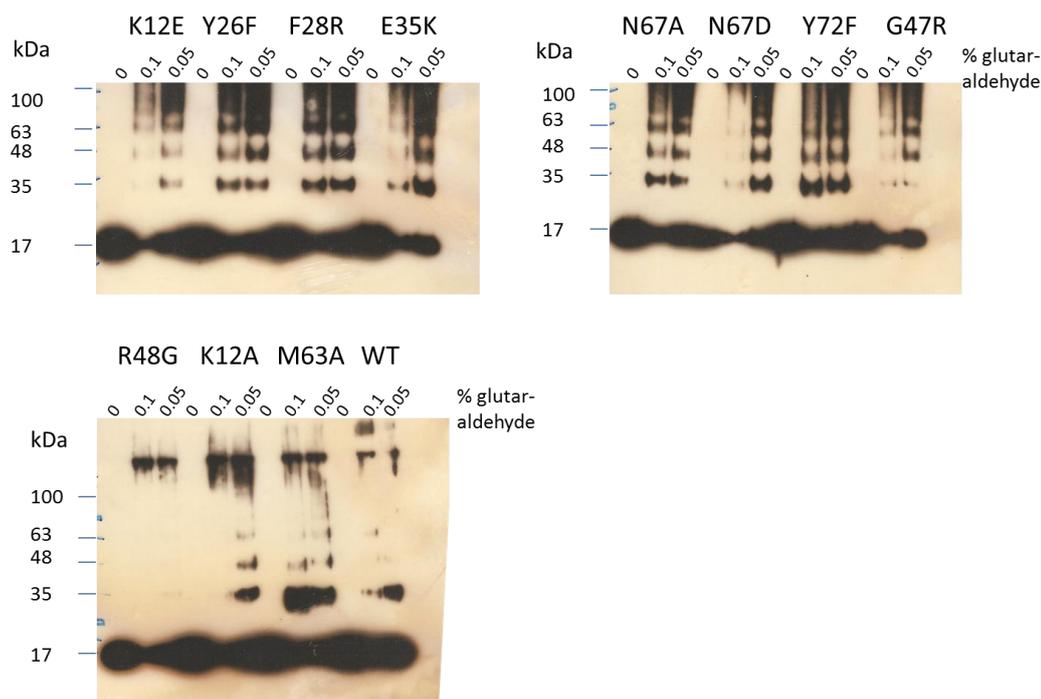
Figure 2.15 Western blots showing the oligomerisation of the Gly1ORF mutant proteins and the wild-type, after treatment with glutaraldehyde.

Glutaraldehyde was used at a final concentration of 0.1% (v/v) and 0.05% (v/v). The aliquots of the reactions were loaded onto 7.5% (v/v) SDS-PAGE gel next to the control with no added glutaraldehyde. The semi-dry transfer was extended to 90 minutes. Proteins were detected using rabbit anti-Gly1ORF1 polyclonal antibodies.

2.2.9 Spectrophotometric analysis of Gly1ORF1 interactions with haem.

Previous work in the Sayers' laboratory demonstrated that Gly1ORF1 forms an insoluble precipitate upon binding haem. The precipitation of Gly1ORF1-haem complexes was the main obstacle in co-crystallisation trials. Two of the generated mutants: Gly1K12A and Gly1M63A, demonstrated much higher solubility compared to the wild type protein, with no precipitant observed when up to 15 μM of protein was incubated with 5 μM haemin (Fig. 2.16).
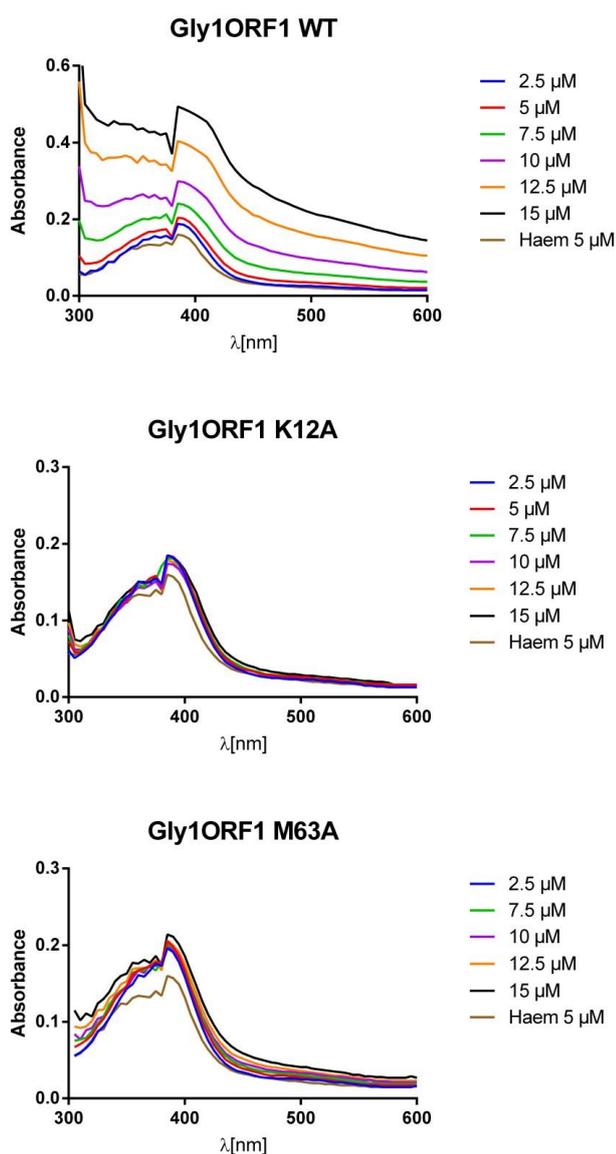
Figure 2.16 Haemin spectra upon binding, with the WT Gly1ORF1 and two mutants demonstrating precipitation.

Increasing concentrations of the WT Gly1ORF1 protein showed increasing levels of precipitation. In the absorbance spectra the precipitant causes a significant increase in the absorbance without changing the maximum absorbance. The Gly1M63A mutant showed some levels of precipitation which was manifested as some distortion in the spectra. The Gly1K12A complexes with haem demonstrated the highest solubility, with the spectra appearing similar for all the tested protein concentrations.

Change in the maximum absorbance of haemin is often indicative of the protein–haemin interaction. There was no shift in the maximum absorbance for any of the tested proteins, however when comparing the spectra of the K12A mutant at the lowest (2.5 µM) and the highest (15 µM) protein concentrations, there was a detectable change in the shape of the curve at the longer wavelengths (Fig. 2.17). This indicates a possible protein–haem interaction, which was observed in the haemin-pull down assay. It is possible that the haem concentration used

was too low to cause a shift; however, low yields of the protein and limited solubility of haem at neutral pH were limiting factors in the experimental design.

The remaining mutants were not tested at such high concentrations due to the low yields of purified proteins; however, initial experiments showed the formation of precipitate in the presence of haem (data not shown).

Figure 2.17 Differences in the haemin spectra in the presence of the Gly1ORF1 K12A protein.

The protein was present at two concentrations; 2.5 and 15 µM. Haem, at a concentration of 5 µM, was used.
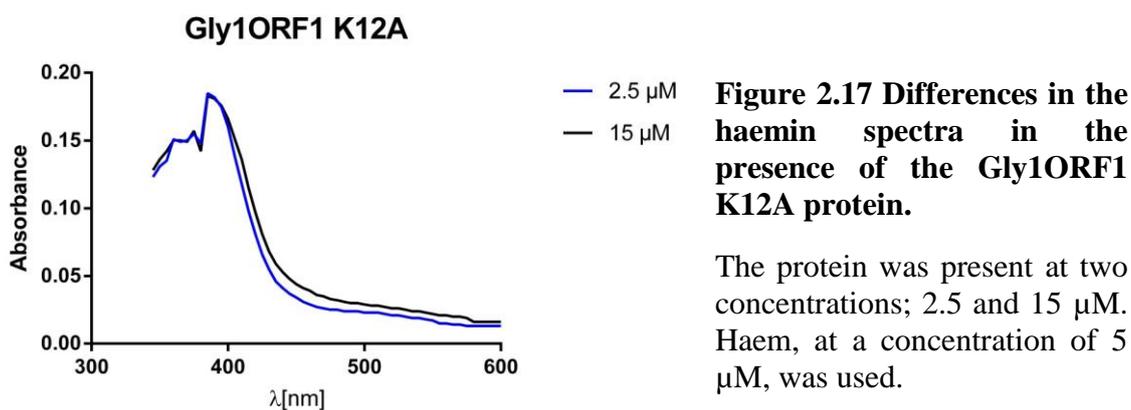
2.2.10 Gly1ORF1 K12A crystallisation trials

The Gly1ORF K12A mutant was selected for crystal trials due to a much higher solubility than the wild-type protein in the presence and absence of haemin. This mutant still bound haemin when tested by the haemin-agarose pull-down assay (See Section 2.1.8). Crystal trials were performed by Jason Wilson, from the Department of Molecular Biology and Biotechnology, University of Sheffield.

The Gly1ORF1 K12A mutant was purified from the culture supernatant as previously described (Materials and Methods, Section 7.6.1), dialysed against a low pH buffer (25 mM NaOAc, pH 4; 50 mM NaCl) and concentrated. The crystals of apo-Gly1ORF1 K12A were grown in 0.2 M NaBr, 0.1 M Bis-Tris propane pH 6.5, and 20% (w/v) PEG3350 using the drop-sitting method, at a protein concentration of 10 mg/ml (Fig. 2.18, Panel A.). The holo-Gly1ORF1 K12A was prepared by incubating the purified protein with haemin solution. The unbound haemin was removed from the solution by small scale gel exclusion chromatography using Zeba Spin Columns (Thermo Scientific).
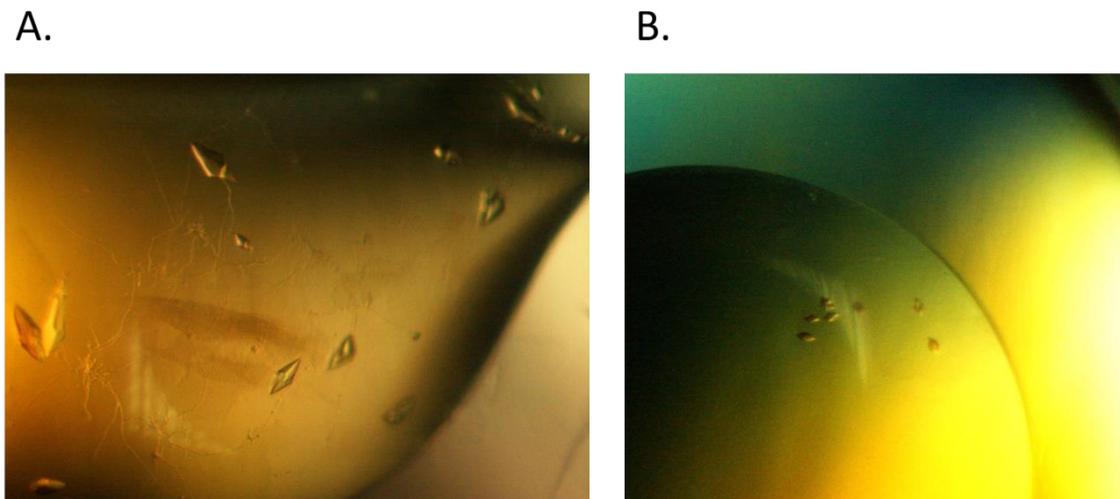
Figure 2.18 Crystals of the Gly1ORF1 K12A mutant.

Protein crystallised without the addition of haemin (**A**) and with haemin (**B**). Pictures kindly provided by Jason Wilson, Department of Molecular Biology and Biotechnology, University of Sheffield.
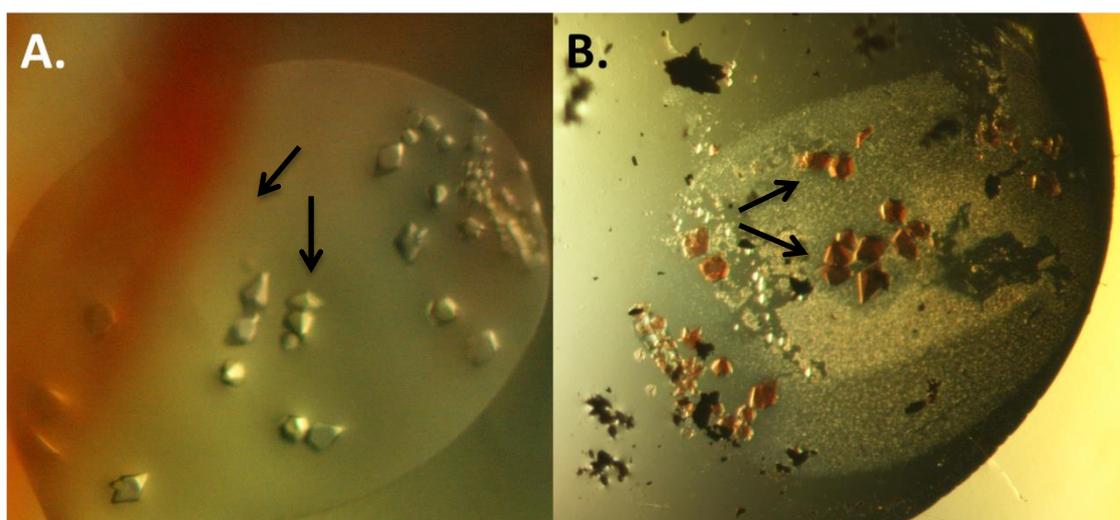
Figure 2.19. Results of haem soaking of the Gly1ORF1 K12A crystals.

A. Crystals of Gly1ORF1 K12A crystals with no haemin; **B.** Crystals after haemin-soaking. Solid haemin was added to the drop containing protein crystals. A change in the colour of crystals to red was observed after several weeks of incubation. Pictures kindly provided by Jason Wilson.

The holo-Gly1ORF1 K12A crystallised in 0.1 M MIB buffer pH 8 and 25% PEG1500 at 2-3 mg/ml (Fig. 2.18, Panel B.). In parallel, crystals of the apo-Gly1ORF1 K12A that crystallised in 0.2 M ammonium acetate, 0.1 M Bis-Tris pH 5.5 and 17% (w/w) PEG 10000 were subjected to crystal soaks. Solid haem crystals were added to the drops containing crystals of the apo-Gly1ORF1 and

incubated for two weeks (Fig. 2.19). Crystals from both the co-crystallisation and crystal soak trials were coloured red, indicating the successful incorporation of haemin into the protein crystals.

The crystal structure of the apo-Gly1ORF1 K12A protein was solved at a resolution of 2.7 Å by Jason Wilson, Department of Molecular Biology & Biotechnology, at The University of Sheffield. The X-ray diffraction for the crystals of the holo-Gly1ORF1 K12A from the haem co-crystallisation trials generated data at a low resolution, which was not sufficient for solving the structure. However, the signal collected by the multichannel analyser (MCA) during data acquisition showed the presence of a peak corresponding to iron, confirming that the co-crystallisation with haemin was successful (Fig. 2.20).
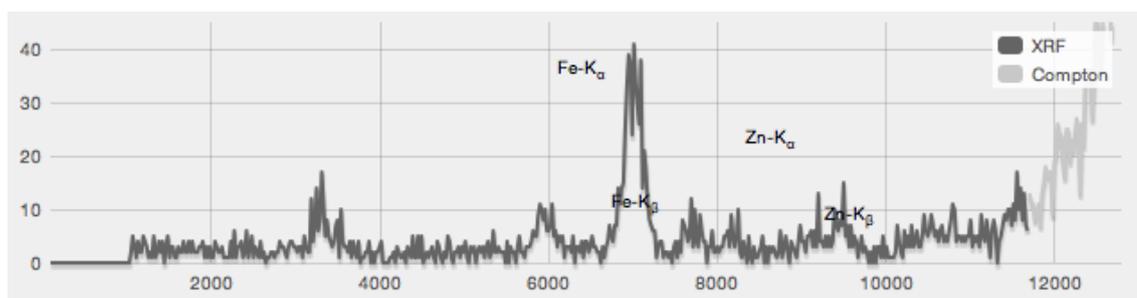
Figure 2.20 A screen shot of the MCA scan of the holo-Gly1ORF1 K12A crystal from the co-crystallisation trial.

2.2.11 Structural analysis of the Gly1ORF1 K12A protein

The crystal structure of the mature recombinant Gly1ORF1 K12A in the apo-form was solved by Jason Wilson (University of Sheffield). The protein, as described previously is present as a dimer in the asymmetric unit (Arvidson *et al.*, 2003). Each monomer consists of two sets of four anti-parallel β-strand sheets arranged orthogonally to each other, and connected by a single α-helix domain (Fig. 2.21).
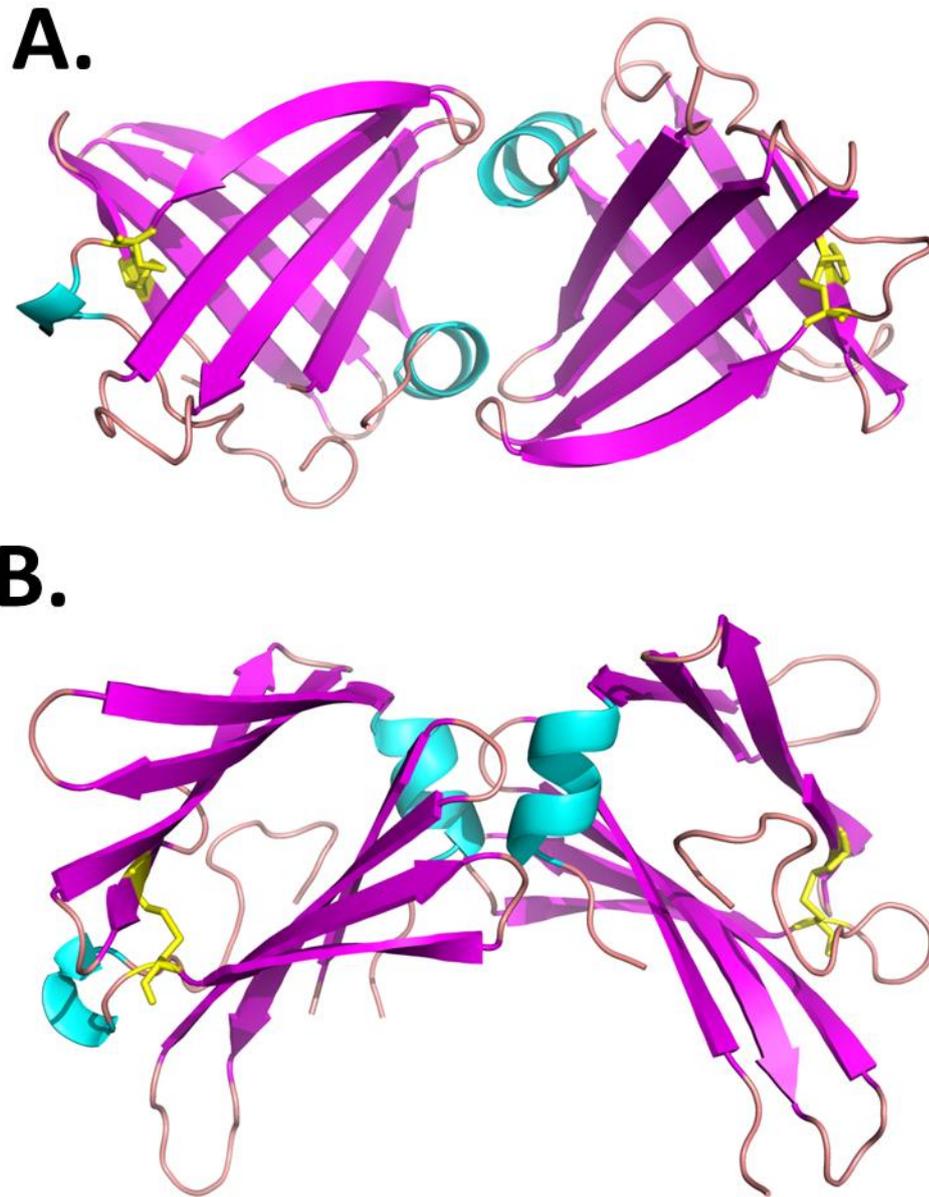
Figure 2.21 A cartoon representation of the Gly1ORF1 K12A structure in two orientations.

The β-strands are shown in magenta, unstructured loops in pink, and α-helices in cyan. The cysteine residues 6 and 103 that form a disulphide bond are represented as yellow sticks. Picture generated in Pymol.

The crystal structure is almost identical to the unpublished structure of the recombinant Gly1ORF1 from *N. gonorrhoeae* MS11, which was solved by Ardvison, and provided to the Sayers laboratory. The replacement of the lysine at position 12 did not cause any obvious changes in the structure that could explain the higher solubility observed for the mutant (data not shown).

The search for structural cavities that could accommodate a haem molecule using the MetaPocket Finder bioinformatics tool (Huang, 2009) revealed the presence of a large pocket at the dimer interface, with the highest Z score 16.02. Several pockets were also identified at different sites; however, with much lower prediction scores (Fig. 2.22).
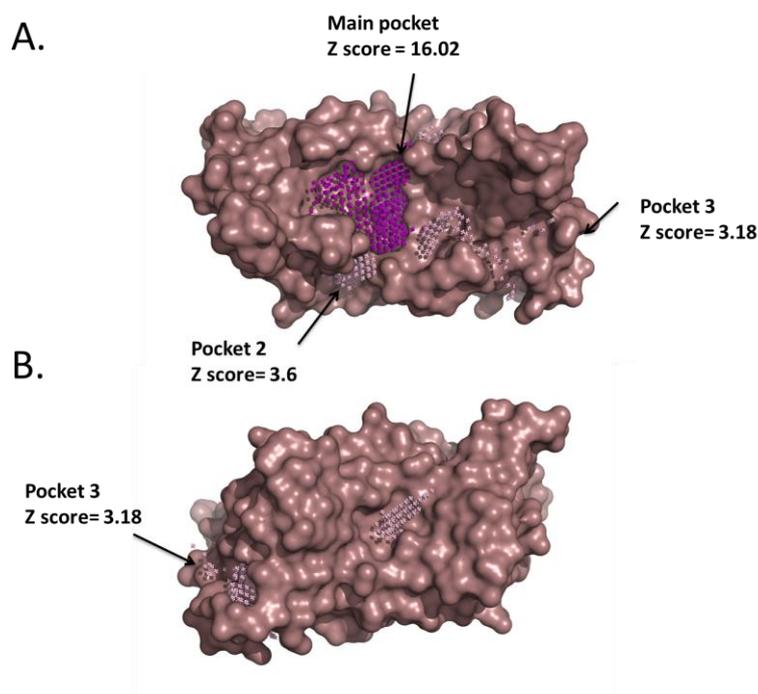
Figure 2.22 Identification of potential haem binding sites in the Gly1ORF1 K12A using the MetaPocket Finder server.

The pocket with the highest score is located in the cavity at the dimer interface, and is depicted as small, purple spheres. Two additional pockets where identified that had a lower score, and are represented as light pink spheres. Images generated and rendered in Pymol.

In order to identify potential haem binding sites on the protein, the docking of haem to the protein was performed using several bioinformatic tools including SwissDock (Grosdidier *et al.*, 2011a, b), Patchdock (Schneidman-Duhovny *et al.*, 2005; Schneidman-Duhovny *et al.*, 2003) and Hex (Macindoe *et al.*, 2010).

The majority of the hits located haem in the largest pocket that was identified by the MetaPocket Finder. The result with the highest prediction score from the haem docking with PatchDock server (Schneidman-Duhovny *et al.*, 2005; Schneidman-Duhovny *et al.*, 2003) is shown in Figure 2.23.
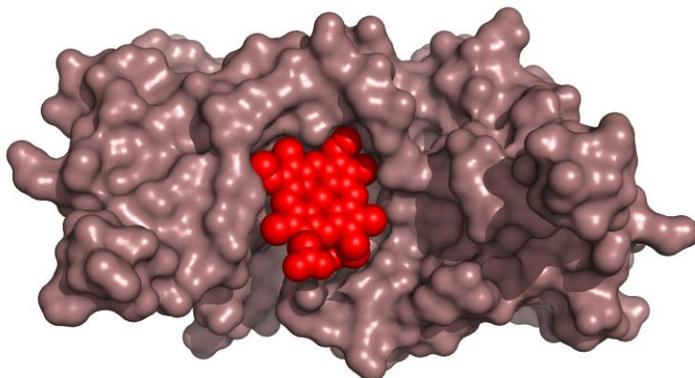
Figure 2.23 Surface representation of the prediction of Gly1ORF1 K12A haem binding site.

The molecular haem docking was done using PatchDock prediction. Haem is represented by red spheres. The image was rendered in Pymol.

Some residues in the Gly1ORF1 K12A structure that are located within 5 Å from the haem molecule in the predicted model (Fig. 2.22) were previously subjected to site directed mutagenesis by Sathyamurthy (Sathyamurthy, 2011). Single amino acid substitutions of aspartic acid 50, lysine 64 and glutamic acid did not inhibit haem binding. In this study, the mutation of arginine 48 to glycine also had no effect on haemin binding and oligomerisation. This may indicate that the changes to the structure or electrostatic potential in Gly1ORF1 were insufficient to disrupt Gly1ORF1 haem interactions. Possibly, mutation of a bigger fragment would result in more dramatic changes in the protein structure and inhibit haem binding. In future studies mutation of the residues 48-50 (shown as yellow sticks in Figure 2.22) could have more prominent effect on the haem binding. These three residues seem to stabilise the haem docking in the cavity at the Gly1ORF1 dimer's interface in the predicted haem binding site.
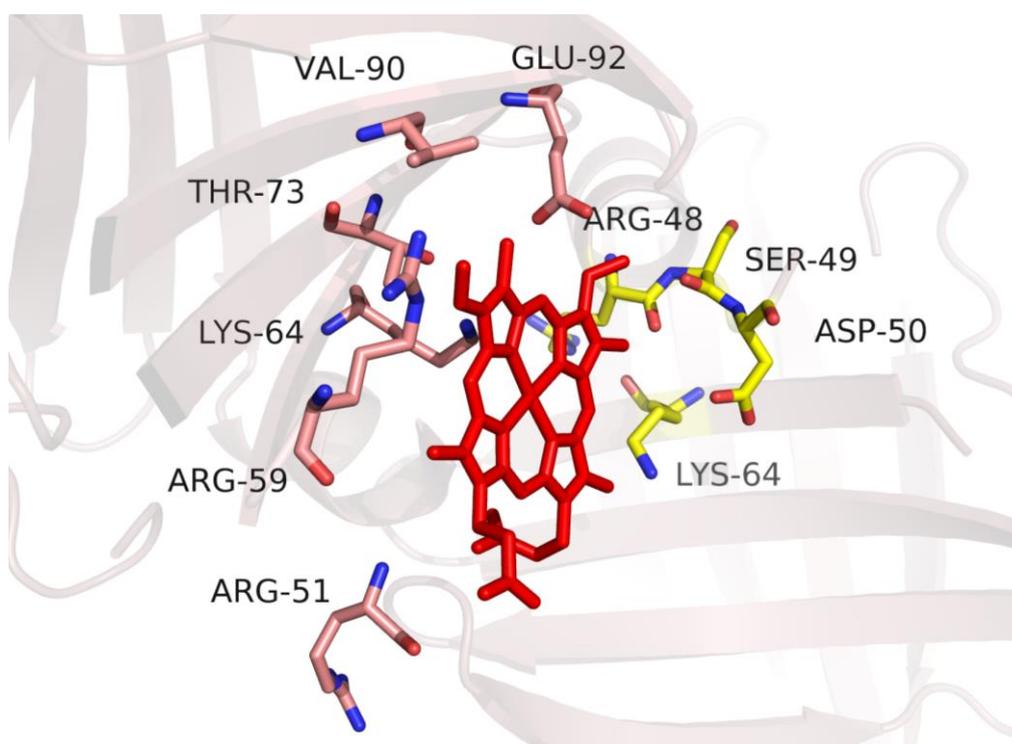
Figure 2.24 Representation of the predicted haem binding site in GlyORF1 K12A structure.

The haem docking was done using PatchDock prediction server (Kelley and Sternberg, 2009). Gly1ORF1 K12A residues within 5 Å of the haem molecule are shown in stick representation. The pink and yellow representation of the residues differentiates the residues from two monomers of Gly1ORF1. Structure of Gly1ORF1 is shown in transparent carton representation. The haem binding site was predicted by MetaPocket finder (Schneidman-Duhovny *et al.*, 2005; Schneidman-Duhovny *et al.*, 2003). The image was generated and rendered in Pymol.

Some residues in the Gly1ORF1 K12A structure that are located within 5 Å from the haem molecule in the predicted model (Fig. 2.24) were previously subjected to site directed mutagenesis by Sathyamurthy (Sathyamurthy, 2011). Single amino acid substitutions of aspartic acid 50, lysine 64 and glutamic acid did not inhibit haem binding. In this study, the mutation of arginine 48 to glycine also had no effect on haemin binding and oligomerisation. This may indicate that the changes to the structure or electrostatic potential in Gly1ORF1 were insufficient to disrupt Gly1ORF1 haem interactions. Possibly, mutation of a bigger fragment would result in more dramatic changes in the protein structure and inhibit haem

binding. In future studies mutation of the residues 48-50 (shown as yellow sticks in Figure 2.24) could have more prominent effect on the haem binding. Those three residues seem to stabilise the haem docking in the cavity at the Gly1ORF1 dimer's interface in the predicted haem binding site.

2.2.12 Binding studies of the Gly1ORF1–haem interaction using microscale thermophoresis

Microscale thermophoresis (MST) is a technique that allows quantitative analysis of protein-ligand interactions. An infrared laser pulse generates a microscale temperature gradient within a small glass capillary. This causes movement of the molecules (thermophoresis), which can be detected by fluorescence. Any changes in the hydration shell, charge or size of the molecule upon binding of a ligand also alters the molecule's thermophoresis. Changes in the thermophoresis or the temperature increase as a function of the ligand concentration, allowing the determination of the dissociation constant (K_d) of the interactions.
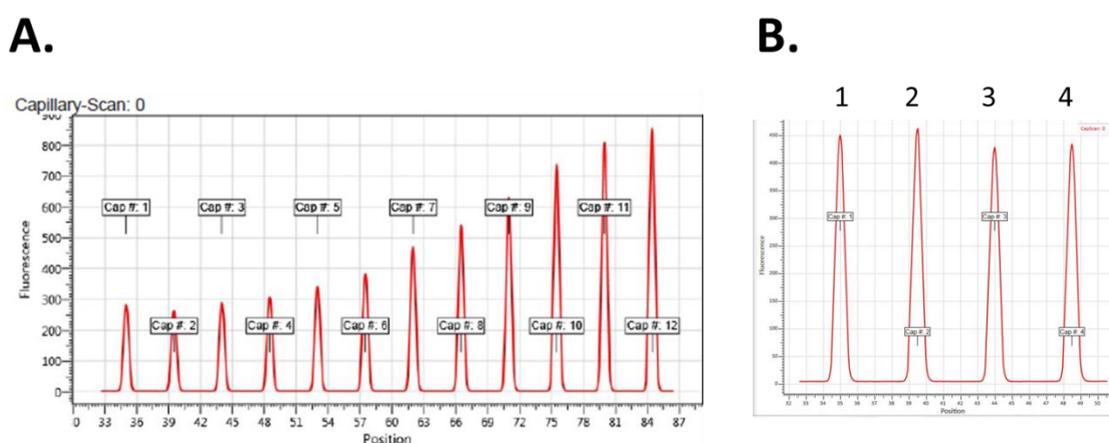
Figure 2.25 Changes in the Gly1ORF1 fluorescence in the presence of haemin.

A. Capillary scan with serial dilutions of haemin, ranging from 50 µM (Capillary 1) to 0.181 µM (Capillary 12). This shows a dose-dependent reduction in the fluorescence of Gly1ORF1 G47R upon the addition of increasing haemin concentrations. The measurements were performed on the Monolith NT.115 Series, at 20% LED power, using the red filter. **B.** The SDS-test: fluorescence scan of denatured Gly1ORF1, with the two highest and two lowest dilutions of haemin (Lane 1 – 0.302 µM, Lane 2 – 0.181 µM, Lane 3 – 50 µM, Lane 4 – 30 µM).

For studying Gly1ORF1 interactions with haemin, affinity purified proteins were fluorescently labelled with Monolith NT^{TM} Protein labelling Kit Red-NHS, according to the manufacturer's instructions. The protein concentration was constant at 200 nM in all tested samples, with the serial dilutions of haemin ranging from 50 µM to 0.181 µM (3:2 serial dilutions). The hydrophilic capillaries were optimal for performing the measurement. The protocol was described in Materials and Methods Section 7.8.3. The experiments were performed on at least two separate occasions with freshly labelled protein.

The determination of the K_d using microscale thermophoresis requires constant concentration of the fluorescently labelled protein in all tested samples. The results of the capillary scan of Gly1ORF1 with serial dilutions of haemin show a dose-dependent reduction in the fluorescence with increasing concentrations of the ligand (Fig. 2.25, Panel A.). In order to determine if the reduction in fluorescence with increased concentration of haemin was a result of protein–haem interaction, or the loss of the protein due to aggregation and precipitation, the SDS test was performed. Mixtures containing a constant concentration of protein and serial dilutions of haemin were prepared as before. The samples with the two highest and two lowest dilutions of the ligand were mixed at a ratio of 1:1 with 2xSD buffer containing 4% (w/v) SDS and 40 mM DTT, and boiled for 5 minutes. The scan showed that upon denaturation, the fluorescence was almost completely restored. This indicates that the reduction in fluorescence with increasing amount of haemin was a result of haemin binding (Fig. 2.25, Panel B.). Therefore, changes of the fluorescence, not the thermophoresis data were used for determining the K_d and EC$_{50}$ for three Gly1ORF1 mutants: G47R, Y72F and M63A. The remaining Gly1ORF1 mutants could not be tested due to the loss of protein during the labelling step and/or problems with aggregation. Analysis of the dissociation constant and plotting the binding curve was performed using the NTAnalysis Software 1.4.27 (Nanotemper Technologies GmbH). Binding curves for Gly1ORF1 G47R mutant are shown in Fig. 2.26).

The data indicate that the protein concentration used for the assay was in the optimal range, as the K_d and EC$_{50}$ values are very similar (Table 2.2). The

concentration of the ligand could not be tested at higher concentrations due to the solubility issue.
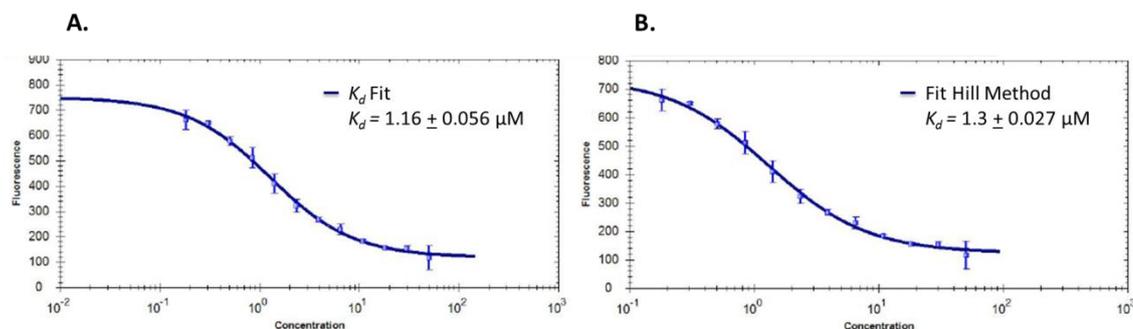
Figure 2.26 Results of the MST analysis of haem binding by Gly1ORF1 G47R.

Haemin was titrated into a constant concentration (0.2 µM) of the labelled protein. Both graphs represent the same data. The graphs represent curves fitted using: **A.** the equation for the K_d determination; **B.** the Hill equation. Error bars show standard deviation of n=3. Results for the Gly1ORF1 variants M63A and Y72F are shown in Table 2.2.

Table 2.2 Affinity haem binding for Gly1ORF1 mutants.

Gly1ORF1 variant	K_d [µM]	EC$_{50}$ [µM]
G47R	1.16 ± 0.056	1.3 ± 0.027
M63A	1.07 ± 0.102	1.38 ± 0.019
Y72F	0.577 ± 0.037	0.892 ± 0.055

The results obtained for the G47R and M63A mutants are very similar. However, the Y72F variant shows almost twice the affinity for haemin compared to the other Gly1ORF1 variants tested. Unfortunately, the affinity of the interaction could not be determined for the wild-type version due to the loss of protein during the labelling step caused by aggregation. It is therefore impossible to say whether Y72F has comparable affinity to the WT and two other mutations caused a reduction in haemin binding, or the opposite: G47R and M63A are comparable to the wild-type, and mutating tyrosine72 resulted in

tighter binding of haem. To answer this question, the haem binding affinity for the wild type Gly1ORF1 must be determined.

2.2.13 Expression of the native Gly1ORF1 using the pET21 expression vector

Although the crystal structure of the Gly1ORF1 K12A with a C-terminal histidine-tag was solved in this study, there is a possibility that the native Gly1ORF1 protein could exhibit different properties by forming dimers with the different interface. Also, having an untagged version of the protein enriches the repertoire of methods that can be used for studying Gly1ORF1 interactions with other recombinant proteins that are expressed with a histidine-tag. Previous attempts to express and purify large amounts of the native protein were unsuccessful. Expression of the native Gly1ORF1 protein in pJONEX4 was very low, and on many occasions was detectable only by western blot (data not shown). This made purification of the native protein extremely difficult.
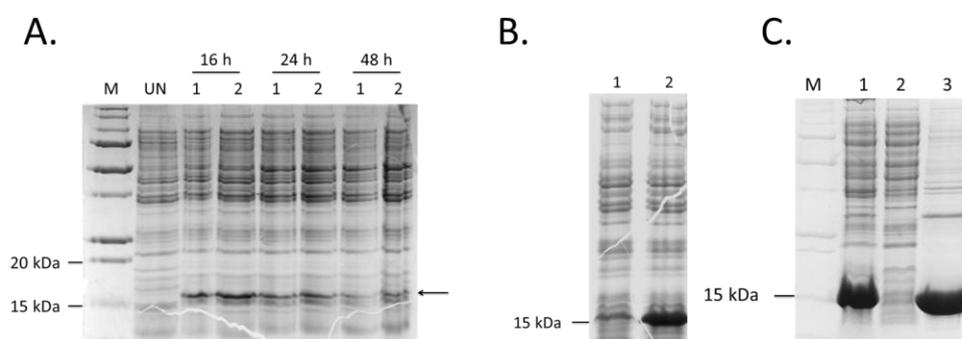
Figure 2.27 Comparing expression of native Gly1ORF1 in pET21 expression vector with (Panel A.) and without (Panel B) the signal peptide.

A. Expression of the full length, native Gly1ORF1 in 1 l auto-induction media. Samples were taken 16, 24 and 48 hours after inoculating. Two 1 litre flasks were inoculated in parallel. **Lane M** – protein marker, **Lane UN** – sample of the inoculum before adding to the auto-induction media; **Lanes 1, 2** – parallel cultures. **B.** Auto-induction of the Gly1ΔSP construct in the pET21 vector**. Lane 1** – sample used as an inoculum, **Lane 2** – sample taken after 24 hours post inoculum. **C.** Cell lysis and sodium deoxycholate wash of inclusion bodies from the 1 litre auto-induction culture. **Lane 1** – Total cell lysate of Gly1ΔSP, **Lane 2** – soluble fraction after centrifugation of the cell lysate, **Lane 3** – inclusion bodies after the sodium deoxycholate wash.

An attempt to produce large quantities of the native protein was performed using the T7-expression system. Sub-cloning the *gly1* gene into the pET21 vector increased the expression levels, but protein yield was still relatively low compared to other Gly1ORF1 variants expressed in this system (Fig. 2.27, Panel A.). To avoid problems with unprocessed signal peptide which could later interfere with the assays, the DNA fragment encoding the signal peptide sequence was removed from the pET21-Gly1ORF1 construct with a single-step PCR method; described in the Material and Methods Section 7.4.2. The expression was tested in 5 ml auto-inducing media. The auto-inducing media contain lactose and glucose as a carbon source. At low cell density, concentration of glucose in the media is higher than lactose and genes under control of *lac* operon are repressed. During bacterial growth, the concentration of glucose in the media decreases. Reduction of glucose concentration below levels of lactose induces expression of those genes.

The expression was successful, and the protein yield was much higher when compared to the culture of the full length *gly1ORF1* gene (Fig 2.27, Panel B.). A large scale auto-induction (see Materials and Methods Section 7.5.4) resulted in even higher yields than the small scale culture. Analysis of the fractions after cell lysis showed that the Gly1ΔSP protein (which lacks the signal peptide) is totally insoluble (Fig. 2.27, Panel C, Lanes 1-2), so a sodium deoxycholate wash of the insoluble fraction was performed as described in Materials and Methods Section 7.6.3.1. The purified inclusion bodies yielded relatively pure protein, and were used for refolding trials (Fig. 2.27, Panel C, Lane 3).

Analysis of the culture supernatants by SDS-PAGE gel failed to detect the presence of Gly1ΔSP and the full-length Gly1ORF1...

2.2.14 Refolding of native Gly1ORF1ΔSP

The use of the Thermo Scientific Pierce Protein Refolding kit allowed the screening of numerous conditions for Gly1ΔSP refolding (see Materials and Methods Section 7.6.3.1). The primary screen with low and high protein concentrations tested the effect of different concentrations of guanidine hydrochloride and L-arginine, as well as a reducing environment (GSH:GSSG)

on protein refolding. After protein refolding in the tested conditions, protein aggregation was measured spectrophotometrically at 450 nm.

Table 2.3 Results of refolding of the native Gly1ΔSP protein using Pierce Protein Refolding Kit.*

Trial/Base Refolding Buffer	Guanidine Hydrochloride [M]	L-Arginine [M]	Redox Environment	Protein precipitation
1	0.4	0	2 mM GSH : 0.2 mM GSSG	Yes – high
2	0.4	0.4	2 mM GSH: 0.4 mM GSSG	No
3	0.4	0.8	1 mM GSH : 1 mM GSSG	No
4	0.9	0	2 mM GSH : 0.4 mM GSSG	Yes - low
5	0.9	0.4	1 mM GSH: 1 mM GSSG	No
6	0.9	0.8	2 mM GSH : 0.2 mM GSSG	No
7	1.4	0	1 mM GSH : 1 mM GSSG	No
8	1.4	0.4	2 mM GSH : 0.2 mM GSSG	No
9	1.4	0.8	2 mM GSH : 0.4 mM GSSG	No

*Final protein concentration – approximately 1 mg/ml

Refolding of low concentrations of Gly1ΔSP (final concentration of approximately 100 µg/ml) in all the tested conditions resulted in no protein aggregation. In trials with high protein concentration (with final concentrations of protein at approximately 1 mg/ml), aggregates were observed in samples with no arginine and low concentrations of guanidine hydrochloride (Table 2.3, Trial 1 and 4), which indicated that the presence of arginine is crucial for Gly1ΔSP refolding. Arginine is known for its ability to supress protein aggregation, although the mechanism remains uncharacterised (Tsumoto *et al.*, 2004). The presence of a reducing environment did not seem to have an effect on aggregation. The composition of each screening condition and the results of aggregation for the high concentration of protein are shown in Table 2.3.
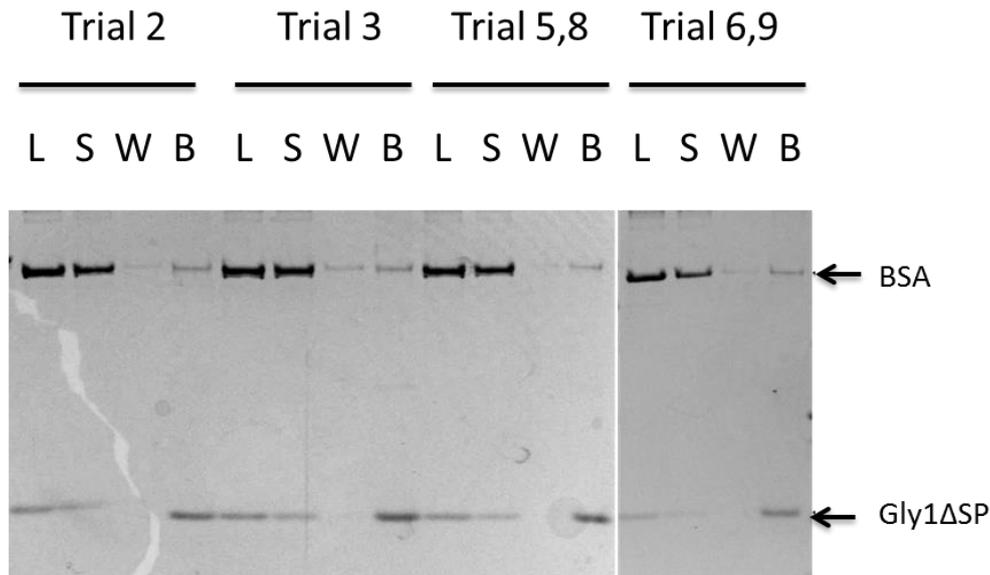
Figure 2.28 Haemin agarose pull down of native Gly1ΔSP after refolding trials.

Aliquots of refolded and dialysed in PBS Gly1ΔSP protein containing BSA as a negative control (**Lane L**) were incubated with haemin agarose beads The beads were centrifuged down and supernatant containing unbound protein was removed (**Lane S**); beads were washed with PBS containing 0.5 M NaCl (**Lane W**) and proteins were eluted from the beads by boiling in SDS-PAGE loading buffer (**Lane B**).

To test whether the protein had correctly refolded, a haemin agarose pull-down assay was performed. Trials 2, 3 were tested separately, whilst trials 5 and 8 were pooled together due to the same concentration of L-arginine. Trials 6 and 9 were pooled together for the same reason. Samples with no arginine in the refolding buffer were not considered for further assays. As the results show, all the tested samples were able to selectively bind to haemin agarose beads (Fig. 2.28).

Size exclusion chromatography (SEC) separated the protein of interest from contaminants that differed in molecular weight. SEC can also be employed as an analytical tool for determining protein size, and for multimerisation studies. In the native Gly1ΔSP refolding and purification process, SEC can help monitor the state of the refolded protein. Correctly refolded protein should give sharp elution peaks. However, if the protein is not properly refolded it elutes at a higher molecular weight as a result of aggregation, or, if the peaks are broad and distorted, the protein has refolded incorrectly.

2.2.14.1 Large scale refolding and size exclusion chromatography of native Gly1ΔSP

Approximately 60 mg of purified inclusion bodies of Gly1ΔSP was refolded in refolding buffer number 3 from the Pierce protein refolding kit, and samples after concentration were loaded onto the size exclusion column (see Materials and Methods Section 7.6.3.1). The elution profile is shown in Fig 2.29.
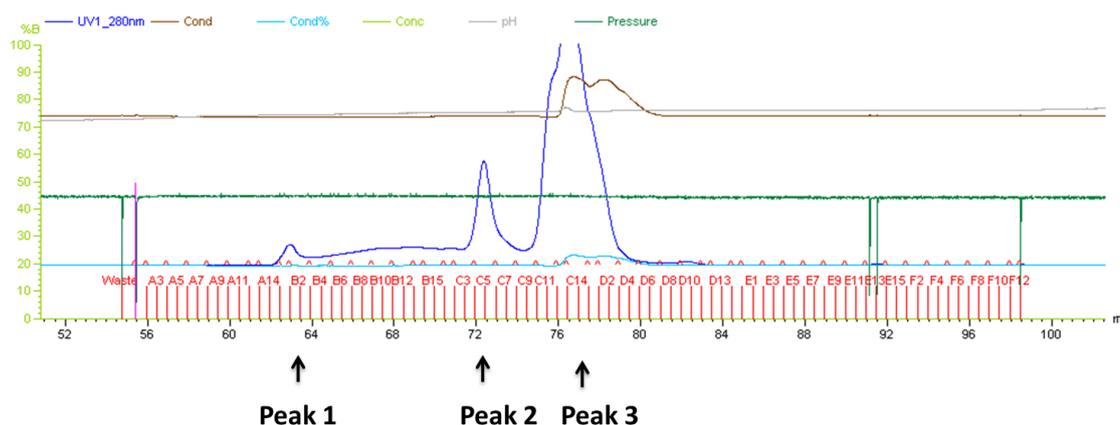
Figure 2.29 Refolded Gly1ΔSP elution from the size exclusion column.

The sample (0.5 ml), at a concentration of approximately 30 mg/ml, was loaded on to a GE Healthcare Superose 6 10/300 GL column and eluted in 0.5 ml fractions. The UV traces of the eluted protein are shown as the blue line, the brown line shows changes in conductivity, and the green line shows the pressure.

Increasing the amount of protein in the refolding conditions did not trigger aggregation on the column, which confirms the choice of refolding buffer. However, some precipitation was observed in fraction C12, which contains the highest concentration of protein (according to the UV-Vis traces). This indicates that the solubility of the protein at very high concentrations must be limited in this buffer.

Three peaks were observed during elution from the SEC column. Fractions corresponding to each peak were pooled together and concentrated ten-fold using the Viva-spin columns and analysed using SDS-PAGE in both non-reducing and reducing conditions. When analysing pooled fractions from the

first peak (A13-B3) by SDS-PAGE, a faint band at the size corresponding to the monomeric Gly1 was visible only in the reducing conditions (Fig. 2.30, Peak 1).
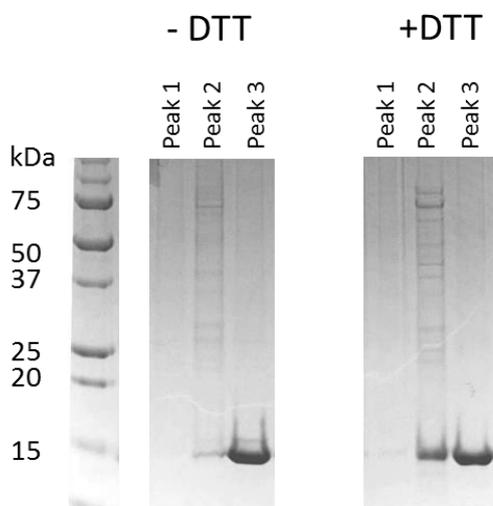
Figure 2.30 Analysis by SDS PAGE of the pooled and concentrated fractions of Gly1ΔSP collected from SEC

Refolded Gly1ΔSP was separated on the GE Healthcare Superose 6 10/300 GL column. Collected samples were analysed by SDS-PAGE in both non-reducing and reducing conditions. **Peak 1** – pooled and concentrated fractions A13 to B4, **Peak 2** – fractions B13-C8, **Peak 3** – fractions C14 to D3, which correspond to the highest peak on the UV-Vis elution profile.

Fractions corresponding to the second peak should contain the dimeric form of Gly1ΔSP. Analysis of the fractions in both reducing and non-reducing conditions did not show any obvious protein at the dimer size; however, there was a difference in band intensity, showing a higher concentration of monomeric Gly1ΔSP in reducing conditions compared to samples run without DTT. Peak 3 corresponds to monomeric Gly1ΔSP and shows that this form is prevalent under these conditions. The presence of an extra band just above the Gly1ΔSP monomer in the non-reducing conditions might indicate that some proportion of the protein was not properly refolded, which affected migration in the SDS PAGE gel (Fig. 2.30).

Changing the buffer to PBS or to a 50 mM Tris-HCl, 100 mM NaCl buffer resulted in precipitation of the protein, indicating that L-arginine is necessary for maintaining protein solubility.

In summary, the Gly1ORF1 protein posed several challenges in the experimental design. Gly1ORF1 is characterised by low solubility in neutral conditions (not exceeding 10 µM) without the addition of stabilising reagents, such as BSA. The protein can be stored at higher concentrations only at low pH, which is far from the physiological conditions. The acidic conditions are also

not compatible for studying protein-haem interactions, as haem solubility is greatly reduced in these conditions. In addition, the protein tends to form large insoluble aggregates upon haem binding, which also limits the number of experimental methods available for studying the protein-ligand interaction. In this study, the introduction of the point mutation (K12A) significantly increased the solubility of the Gly1ORF1 protein. The protein retained its haem binding activity without the formation of insoluble aggregates. This enabled the successful crystallisation of the protein, which allowed the crystal structure of the haem-free protein to be determined. Co-crystallisation trials with haemin also gave promising results. However, the data collected for the crystals of the holo-Gly1ORF1 K12A were too low resolution to be solved. Therefore, the conditions for co-crystallisation of Gly1ORF1 K12A with haemin must be further optimised.

Another problem encountered in this project was inconsistent expression of the protein in the heat-inducible expression system. The yields of the protein present in the cell pellets were low, and the protein was unprocessed and insoluble. However, the protein was also secreted in the culture supernatants. The yields and purity of the protein differed from batch to batch. In this study, a gene encoding the Gly1ORF1 native protein without a signal peptide (Gly1ΔSP) was cloned into T7-expression vector. This resulted in a significant increase in protein yields in the cell pellets. However, the protein was present in inclusion bodies. Nevertheless, the levels of protein expression in the T7-inducible system proved to be reproducible and easily controlled when compared to the heat-inducible system. For future experiments requiring high yields of protein, mutated *gly1ORF1* genes could be sub-cloned into the same T7-expression system without the signal peptide.

Problems in purifying some Gly1ORF1 mutants from the culture supernatants by IMAC (Ion-Metal Affinity Chromatography) were partially resolved by enriching the media with $NiCl_2$, or by ammonium sulphate precipitation. However, it required handling large volumes of supernatant, and the yields depended also on the efficiency of protein secretion, which was not reproducible and varied between mutant proteins. Therefore, producing the protein as inclusion bodies at high yields, and refolding it in the conditions

optimised in this study seems to be a better alternative. However, there is always a risk that even a single point mutation could have an impact on the refolding process, and each mutant would require separate optimisation of the refolding conditions. This would be time consuming.

There is still a need for optimising the buffer conditions for protein storage at high concentrations, in neutral conditions without the use of interfering additives. The use of agents, including L-arginine, which seems to increase the protein solubility at pH 7-8, does interfere with labelling reactions. BSA, which stabilises the protein, is also able to bind haem, and therefore cannot be used in haem binding experiments. There is an increasing number of commercial kits that allow screening of a wide range of buffer conditions in order to optimise a particular protein's solubility. It is possible that the addition of a low concentration of detergent or inert solvent, such as DMSO could be beneficial for protein stability and preventing aggregation. Such measures would greatly facilitate studies on the Gly1ORF1 protein.

CHAPTER 3   STUDIES ON GLY1ORF1 INTERACTIONS

3.1 Introduction

Gly1ORF1 was first identified as a protein that confers a weak haemolytic activity when expressed in *E. coli* (Arvidson *et al.*, 1999). Further experiments in our laboratory demonstrated that Gly1ORF1 is able to bind to red blood cells when analysed by FACS and causes morphological changes of red blood cells characterised by clumping and loss of the discoid shape (Sathyamurthy, 2011). The yeast two-hybrid experiments by Meadows showed that Gly1ORF1 specifically binds membrane protein EPB4.2 – one of the main components of the red blood cell cytoskeleton (Meadows, 2004), which was further confirmed by Sathyamurthy. The experimental data showed interactions with mammalian cell lines expressing recombinant EPB4.2 protein (Sathyamurthy, 2011). The fact that EPB4.2 is not expressed on the surface of the cells, but is present on the cytoplasmic face of the membrane prompts the question of how Gly1ORF1 enters the cells to allow the interactions with its identified binding partner. It was hypothesised that Gly1ORF1 might have a membrane receptor on the surface of the red blood cells allowing it to be translocated across the cell membrane. In this chapter the aim was to investigate whether Gly1ORF1 interacts with other proteins on the red blood cell membrane that could be potential receptors using a cross-linking approach. It was also investigated whether Gly1ORF1 binding to red blood cells is reversible by performing subsequent washes and analysing the amount of the protein that was bound to red blood cells.

Also, the experiments in this chapter describe cloning and expression of EPB4.2 in *E. coli* and purification of the protein. The aim was to co-crystallise EPB4.2 and Gly1ORF1 protein and study the interactions between the proteins using the biolayer interferometry approach.

Gly1ORF1 protein has been shown to bind haemin *in vitro* and the data by Sathyamurthy demonstrated that the Gly1ORF1 knock-out mutant is not able to grow on haem or haemoglobin as the sole iron source (Sathyamurthy, 2011).

This suggested that Gly1ORF1 might be involved in iron acquisition in *Neisseria.* Two haem receptors were identified in *Neisseria:* HmbR and HpuAB, which were described in Section 1.2.6.3. It was hypothesised that Gly1ORF1 might play a role of a haemophore for the neisserial outer membrane receptors by capturing haem from the extracellular environment and transporting it to the outer membrane receptors. Similar haemophores were described in many bacteria, e.g. HasA in *Serratia marcescens* (Czjzek *et al.*, 2007).

A further aim of this chapter was to clone and express neisserial haem receptors and investigate if Gly1ORF1 is able of interact with them.

3.2 Results and Discussion

3.2.1 Interactions with red blood cells

3.2.1.1 Reversibility of binding to red blood cells

This experiment aimed to investigate whether the interactions of Gly1ORF1 with red blood cells are reversible. Human red blood cells from a healthy volunteer were first incubated with Gly1ORF1 and then washed sequentially with PBS for 15 minutes each time by gentle mixing (see Materials and Methods Section 7.9.4). Samples of the red blood cell suspension were taken for SDS-PAGE followed by western blot with anti-Gly1ORF1 antibodies. The immunoblot showed that the amount of Gly1ORF1 bound to red blood cells was significantly reduced after first two washes with no further reduction after the third wash (Fig. 3.1). It is not clear whether the bound Gly1ORF1 traversed the membrane and was captured inside red blood cells or if there is strong interaction with membrane proteins that were not disrupted after sequential washes. Gly1ORF1 tends to precipitate in the presence of haemin and it might also be possible that some lysis occurred during incubation and free haem interacted with Gly1ORF1 causing precipitation. Gly1ORF1 could then be pelleted during centrifugation together with intact red blood cells and was not removed after aspiration of the wash buffer.

It is quite interesting that even after addition of DTT in the loading buffer Gly1ORF1 is still present in multimeric forms. It is possible that upon binding to the target protein on the red blood cells, complexes are very stable and can have limited accessibility for the reducing reagent or that components of the red blood cells such as haemoglobin could partially interfere with the DTT reaction as previously reported (Netto and Stadtman, 1996).
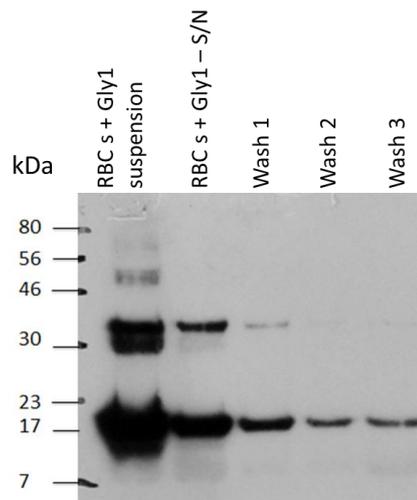
Figure 3.1 Immunoblot showing reversible binding of Gly1ORF1 to human red blood cells.

Red blood cells (RBCs) were incubated with Gly1ORF1. The RBCs were centrifuged down and supernatant (S/N) with unbound protein was removed. RBCs were subsequently washed with PBS. A 10 µl aliquot of red blood cell suspension was taken after each wash. The aliquots of red blood cells with bound proteins after each wash were analysed by western blot with anti-Gly1ORF1 antibodies.

3.2.1.2 Cross-linking of Gly1ORF1 with red blood cells

It is still not clear whether Gly1ORF1 has a receptor in the membrane of red blood cells allowing translocation of the protein inside the cells and interaction with protein EPB4.2. To try to identify potential receptor in red blood cells, Gly1ORF1 was incubated with red blood cell suspension and two reversible cross-linkers were added (see Materials and Methods Section 7.9.3). DTSSP (3,3'-dithiobis[sulfosuccinimidylproprionate]) is reversible cross-linker that is not membrane-permeable and can cross-link only cell surface proteins, whereas DTBP (dimethyl 3,3'-dithiobispropionimidate) is able to cross the cell membrane and target proteins inside the cells. The advantage of using those cross-linkers is reversibility of the reaction. Addition of DTT to the sample cleaves the spacer arms in coupled proteins. Red blood cells after cross-linking were lysed and analysed by western blot with anti-Gly1ORF1 antibodies in both non-reducing and reducing environments. The results of the cross-linking with both reagents

did not show any obvious proteins at high molecular weight indicating interactions of Gly1ORF1 with RBC membrane proteins (Figure 3.2). Although interactions with EPB4.2 were previously shown in our lab, no bands corresponding to the complex of EPB4.2 with Gly1ORF1 size (approximately 88 kDa) were detected. Interestingly, some extra bands running at a lower molecular weight than Gly1ORF1 monomer and dimer were detected in non-reducing gels which were not present in the reducing conditions. The band just below 30 kDa could be Gly1ORF1 cross-linked to some other small protein. It could be also a more compact form of a Gly1ORF1 dimer that migrates faster in the SDS-PAGE gel, which needs to be further investigated.
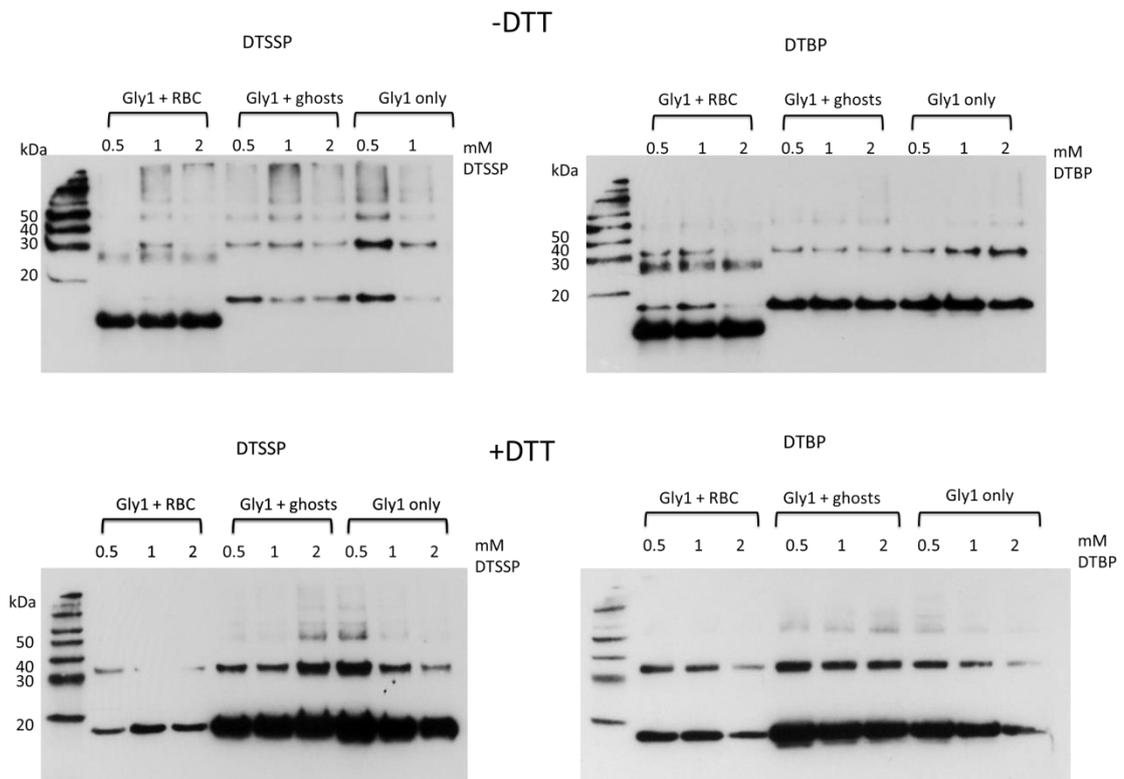
Figure 3.2 Immunoblot of Gly1ORF1 cross-linking with red blood cells.

Gly1ORF1 was incubated with red blood cells or ghosts in suspension and subjected to reversible cross-linking with DTSSP and DTBP. Gly1ORF1 with no added cells or membranes was used as a control. After cross-linking, red blood cells and ghosts were centrifuged down, lysed and run on SDS PAGE and analysed by western blot with anti-Gly1ORF1 antibody. Samples were run in non-reducing (-DTT) and reducing conditions (+DTT). The Gly1ORF1 was detected using anti-Gly1ORF1 polyclonal antibodies.

It is not clear why a monomer of Gly1ORF1 after treatment with both cross-linkers showed a different gel migration pattern in the presence of red blood cells compared to Gly1ORF1 incubated with ghosts or without red blood cells and membranes. One of the explanations is conformational changes in the Gly1ORF1 structure such as domain swapping triggered by the presence of red blood cells and fixed by the presence of cross-linkers. If the changes were triggered by the presence of the red blood cell membranes, samples subjected to cross-linking with ghosts would be expected to show similar pattern. One of the explanations of such changes could be the presence of haem. A question of how Gly1ORF1 gets inside red blood cells remains unanswered.

3.2.2 Gly1ORF1 interactions with recombinant EPB4.2

3.2.2.1 Sub-cloning and expression of membrane protein EPB4.2

EPB4.2 was codon optimised for expression in *E. coli* and the gene was synthesized by Eurofins with a sequence encoding the N-terminal histidine tag. The gene was then subcloned from the pUC57 vector into the pET21 vector using NdeI and HindIII restriction sites. The construct was propagated in *E. coli* XL1Blue and expressed in *E. coli* BL21DE3 cells by induction with 0.1 mM IPTG (see Materials and Methods Section 7.5.3),

Analysis of induced cell lysates by SDS-PAGE analysis showed a detectable band at approximately 72 kDa in the induced samples (Fig. 3.3, Panel A). The expression levels of EPB4.2 were much lower than the positive control with N-biotin tag Gly1ORF1 (Fig. 3.3, Panel A, Lane cntr). Western blot analysis of the induced samples showed some protein degradation after 4 hour-induction, indicating that the protein is unstable (Fig. 3.3, Panel B). Thus, for large scale expression, the protein induction was limited to two hours. An attempt to purify the protein by Ni--affinity chromatography under native conditions was unsuccessful due to protein degradation even after adding EDTA-free protease inhibitors cocktail (Roche) to solutions used for purification. This indicates that the protein might be susceptible to proteases that are not inhibited by the EDTA-free reagent. However, EDTA is not compatible with the Ni--affinity column and had to be avoided. Therefore, the protein was purified in denaturing conditions using guanidine hydrochloride (see Materials and Methods Section

7.6.3). Protein was bound to the column in the presence of 3 M guanidine hydrochloride and was refolded on the column. Fractions containing N-histidine tag EPB4.2 were concentrated, buffer exchanged into PBS and used for studying protein-protein interactions (Fig. 3.4, Panel A.). The western blot analysis with the mouse monoclonal anti-EPB4.2 confirmed that the purified protein is EPB4.2 (Fig. 3.4, Panel B).
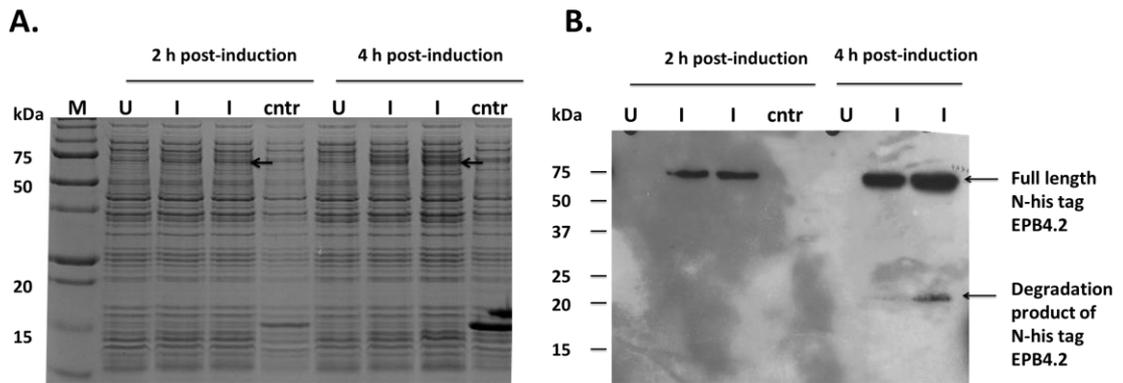
Figure 3.3 Analysis of expression of N-histidine tag EPB4.2 in BL21DE3.

Analysis performed: by SDS-PAGE (Panel A) and by western blot (Panel B). Mid-log phase duplicate cultures were induced with 0.1 mM IPTG and samples 2 and 4 hours were run on the SDS-PAGE gel and subjected to western blot with anti-histidine tag mouse mAbs. The presence of a faint band at approximately 72 kDa size was detected by SDS-PAGE (indicated with arrows). Western blot analysis showed an extra band of approximately 22 kDa which is a degradation product of the expressed protein. Culture of Bl21DE3-pET21(N-biotin-Gly1ORF1) was used as a positive control for the induction conditions (**Lane cntr**).
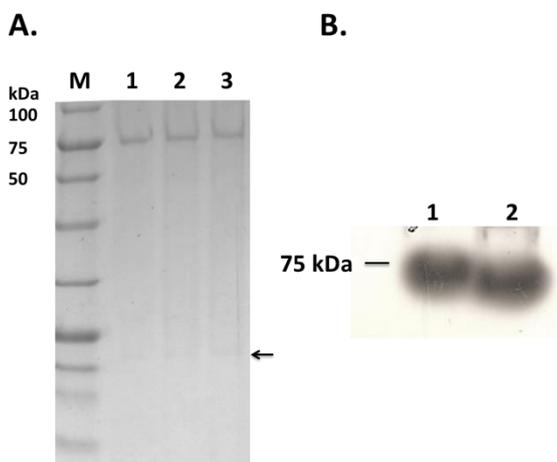
Figure 3.4 EPB4.2 purification on Ni-affinity column under denaturing conditions.

A. SDS-PAGE gel showing ten times concentrated fractions (3-5) (Lanes 1 to 3) collected from Ni-affinity column in linear gradient of imidazole. Samples of 5 μl from each fraction were loaded in each lane. The band corresponding to the degradation product at 22 kDa is marked with an arrow. **B.** Western blot analysis of the purified recombinant protein with the mouse monoclonal anti-EPB4.2 antibody (Abcam). **Lane 1** – purified recombinant N-histidine tagged EPB4.2; **Lane 2** – positive control for the antibody – isolated red blood cells membranes (2 μg in the lane).

The yields of the purified protein were low (less than 0.3 mg from 12 g of induced cells); however, the amount was sufficient to perform an initial test for detection of the interaction with Gly1ORF1. In order to study whether the recombinant N-histidine tag EPB4.2 can bind to Gly1ORF1 *in vitro*, a biolayer interferometry approach was used. A BLITz machine (ForteBio) was used for the study. The method involves measuring changes of the white light interference reflected from the biosensor surface with reference to the internal layer. Binding of the molecules to the surface of the biosensor causes changes in the interference pattern that can be measured in real time. For studying protein-protein interactions one of the binding partners is immobilised on the surface of the biosensor and is incubated with different concentrations of the second partner. The method allows quantitative measurement of the association and dissociation of a ligand from the biosensor, thus allowing determination of binding kinetics of two molecules.

For detecting the interactions between EPB4.2 and Gly1ORF1, N-histidine tag EPB4.2 was first immobilised covalently on AR2G biosensors (Amine-reactive 2nd Generation biosensors) according to the manufacturer's instructions and incubated with the Gly1ORF1 C-histidine tag solution. However, the AR2G sensors do not allow a directional immobilisation of proteins with high numbers of lysine residues. Therefore, a Gly1ORF1 binding site in the large proportion of immobilised EPB4.2 protein molecules might be inaccessible. An alternative - Ni-NTA sensors enable immobilisation of N-histidine tag EPB4.2 in a similar way as the protein is anchored to the red blood cell membrane via N-terminal mirystyl-group (Das *et al.*, 1994). Unfortunately, this approach could not initially be used. The binding partner also contained a histidine-tag and would interact with the sensor. Only at the later stage of the project, success in the purification and refolding of the non-tagged Gly1ORF1 allowed the use of Ni-NTA biosensors. The N-histidine tagged EPB4.2 coupled to Ni-NTA biosensors via its N-histidine tag was incubated with the native Gly1ORF1 to analyse protein-protein interactions.

In both cases no interaction between the proteins was detected (Fig. 3.5). The main reason for that might be improper refolding of the EPB4.2 protein on the Ni-affinity column. The instability of the protein which could be observed even

during protein expression in *E. coli* dictated the choice of the denaturing purification conditions. Possibly, the addition of a protease inhibitor to the cell culture shortly after induction and before harvesting the cells might prevent or decrease the degradation and allow purification in the native conditions. Another approach could include a choice of a different Ni-affinity resin that would tolerate the presence of EDTA in the purification buffers and thus ensure inactivation of the metalloproteases.
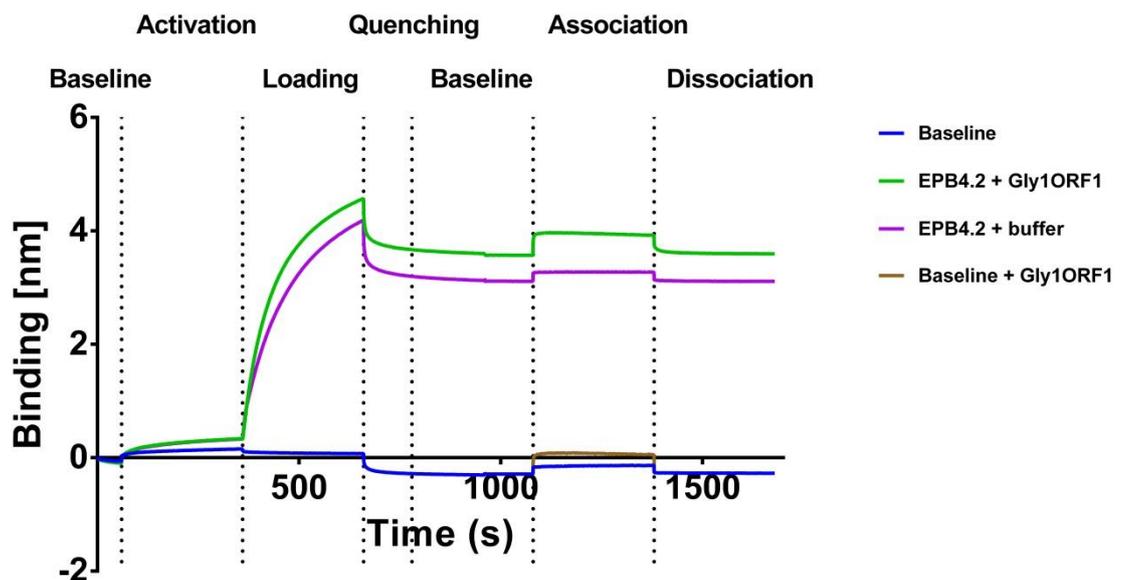
Figure 3.5 EPB4.2 interactions with Gly1ORF1 analysed by biolayer interferometry (Blitz).

EPB4.2 was first immobilised on an AR2G biosensor. After setting up a baseline, the biosensor was incubated with Gly1ORF1 C-histidine tag protein. Controls included a baseline with no protein immobilised incubated with buffer and Gly1ORF1 solution at 100 µg/ml and a biosensor with immobilised EPB4.2 incubated with buffer only. No obvious binding of Gly1ORF1 was observed.

Possibly eukaryotic proteins expressed in *E. coli* might have limited function due to the lack of post-translational modifications. EPB4.2 undergoes myristoylation at the N-terminus to allow the attachment of the protein to the red blood cell membrane. Presence of the N-histidine tag in the recombinant EPB4.2 protein plays an analogous role – it allows attachment to a surface, the biosensor, so it was thought very unlikely that this would affect protein-protein interactions. However it cannot be ruled out that other post-translation modifications are important for the interaction of EPB4.2 with Gly1ORF1.

Other approaches to study *in vitro* Gly1ORF1 and EPB4.2 interactions were tested. Purification of the protein from human and porcine erythrocyte membranes using the protocol described by Malik and co-workers (Malik *et al.*, 1993) was unsuccessful, also probably due to the protein degradation during purification process. Cross-linking experiments and pull-down assays with the detergent treated ghost membranes and C-histidine tagged Gly1ORF1 did not show any obvious interactions (data not shown). Finally, collaboration was set up with Dr Scott Garforth from the Albert Einstein College of Medicine in New York to obtain a purified EPB4.2 expressed in murine cell culture for further experiments and co-crystallisation trials. Unfortunately, although the purification of small quantities of the protein was successful, again, the protein showed a high degree of degradation and was not suitable for further analysis.

3.2.3 Studying interactions of Gly1ORF1 with haemoglobin receptors

3.2.3.1 Cloning and expression of haemoglobin receptors

Oligonucleotides for amplification of the *hmbR* gene were designed based on the *N. meningitidis* MC58 genomic sequence (Accession number: gi77358697:1740259-1742643), whereas oligonucleotides for cloning of *hpuA* and *hpuB* genes were based on sequence of haemoglobin-haptoglobin utilisation operon of *N. meningitidis* (Accession number: gbIU73112.2INMU73112:129-3686). Oligonucleotide sequences and PCR conditions are summarised in Table 7.1 in Materials and Methods Section 7.3.2.

A 2376 bp gene encoding *hmbR* was successfully amplified from NM MC58 genomic DNA (Fig. 3.6). However, no DNA fragments corresponding to *hpuA* and *hpuB* genes were obtained even after several attempts of optimisation. No *hmbR* products were observed for *N. gonorrhoeae* MS11 or *N. lactamica* Y92-1009, which agrees with the results of sequence analysis of genomes of those strains and data available in the literature. The PCR product of *hpuB* gene was successfully amplified from *N. lactamica* Y92-1009 genomic DNA. Although this gene was not identified in the available genome of NL Y92-1009, this genomic sequence is not complete and is probably missing the *hpuAB operon* region as well as the region encoding *gly1ORF1*, which was also amplified with the positive control primer. Surprisingly, the positive control with primers for the

amplification of *gly1ORF1* using genomic DNA of NG MS11 as a template gave a faint band corresponding to the expected 1.3 kbp – long fragment. In addition, a more intense DNA fragment of a larger fragment (approximately 1.8 kbp) was observed on the gel. It is possible that through several passages NG MS11 started to undergo phase variation, which could result in the presence of those two products. It was reported in the literature that *Neisseria* spp*.* can undergo those changes even after relatively few passages (Tauseef *et al.*, 2013).
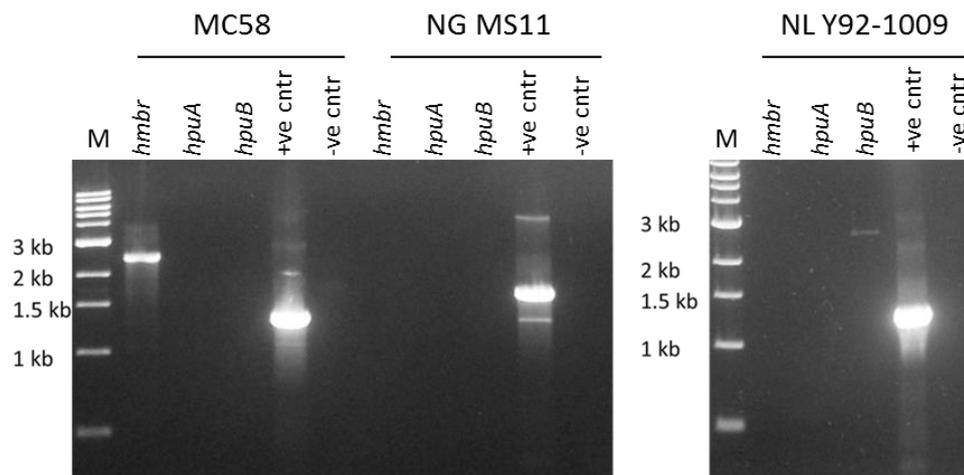
Figure 3.6 Amplificiation of haemoglobin receptors genes from genomic DNA of three strains of *Neisseria* spp.

Genomic DNA of *N. meningitidis* MC58, *N. gonorrhoeae* MS11 and *N. lactamica* Y92-1009 were used as template for PCR of the *hmbr, hpuA* and *hpuB* DNA fragments. Positive control contains amplified *gly1ORF1* DNA fragment (**Lane +ve cntr**). **Lane –ve cntr** - negative control; **Lane M** – size marker;

HpuB cloning was unsuccessful. The *hmbR* gene of MC58 was cloned into the pET21 vector and successfully expressed in BL21Star(DE3) cells. Expression was confirmed by SDS-PAGE gel of induced samples stained with standard Coomassie blue stain (Fig. 3.7, Panel B). A protein of the expected size of approximately 90 kDa was readily visible on the gel. To ensure the produced protein was the haemoglobin receptor, induced cell lysates were subjected to SDS-PAGE separation and incubated in TBG buffer without SDS, allowing partial *in situ* refolding of the proteins (see Materials and Methods section 7.10.2). The gels were then incubated with Congo red or haemin solutions in TBG buffer (Fig. 3.7, Panel A. and 3.7, Panel C). Congo red has been

previously used to identify haem binding protein in bacteria (Daskaleros and Payne, 1987; Smalley *et al.*, 1995; Stugard *et al.*, 1989). It has been used as an alternative to haem due to its higher solubility in water.

Overexpressed recombinant HmbR bound both Congo red and haemin similarly to the Gly1ORF1 used as a control for staining.
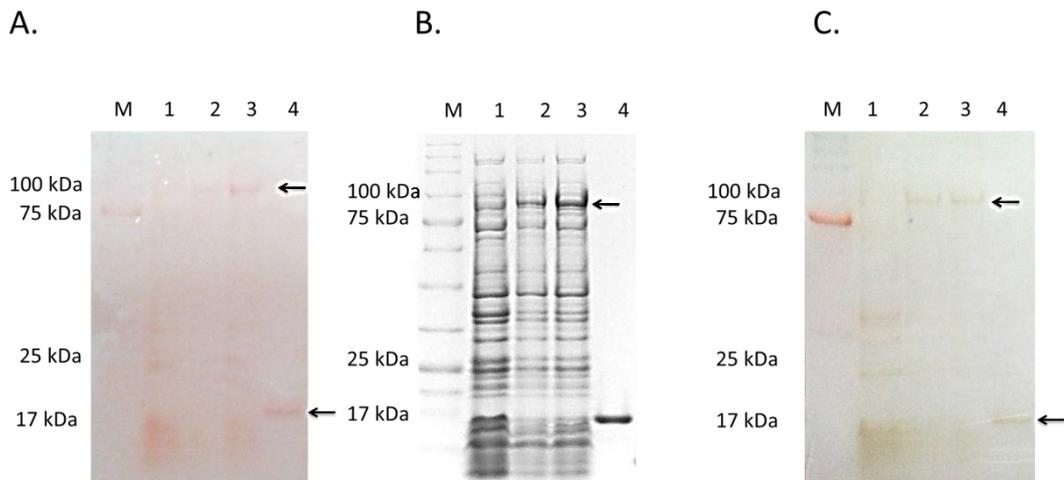
Figure 3.7 Analysis of small scale induction of HmbR protein in *E. coli* BL21Star (DE3)

A. SDS-PAGE gel red after incubation in TBG buffer stained with Congo red. **B.** Standard SDS-PAGE gel stained with Coomassie blue stain **C.** SDS-PAGE gel red after incubation in TBG buffer stained with haemin. Results confirm binding of Congo red and haemin by expressed recombinant HmbR after in-gel refolding. **Lane M** – size marker; **Lane 1** – negative control with a non-inducing HmbR clone, **Lane 2** and **3** – clones with wild-type *hmbR* gene, **Lane 4** – GlyORF1 C-histidine tag as a positive control of staining.

To check whether HmbR is expressed on the surface of the recombinant cells, freshly transformed BL21Star(DE3) grown for two hours in non-inducing liquid medium were streaked out on inducing plates containing selection antibiotic, lactose and Congo red (see Materials and Methods Section 7.10.1). Bacteria harbouring the pET21 vector containing the *hmbR* gene grew as red colonies indicating selective binding of Congo red and confirming expression of the membrane bound receptor. In contrast, colonies of BL21Star(DE3) with an empty pET21 vector were pink with red dye present only in the centre of the colonies. Presence of the halo around the colonies expressing HmbR indicates selective binding of the dye (Fig. 3.8).

It was also tested whether bacteria with expressed HmbR can bind haemin and haemoglobin in solution (see Materials and Methods Section 7.10.3). Bacteria grown in liquid medium and induced with IPTG were tested for the ability to sequester haemin or haemoglobin from the solution. Samples were taken at 30 minute time-points post-induction, the optical density ($OD_{600\ nm}$) of the bacterial suspension was adjusted to 1 and the bacterial suspension in PBS was incubated with haemin or bovine haemoglobin. The percentage of bound haemin and haemoglobin in the samples was determined spectrophotometrically by measuring the reduction in absorbance of the supernatants after pelleting the cells. Samples with haem or haemoglobin only were used as a reference (Fig. 3.9).
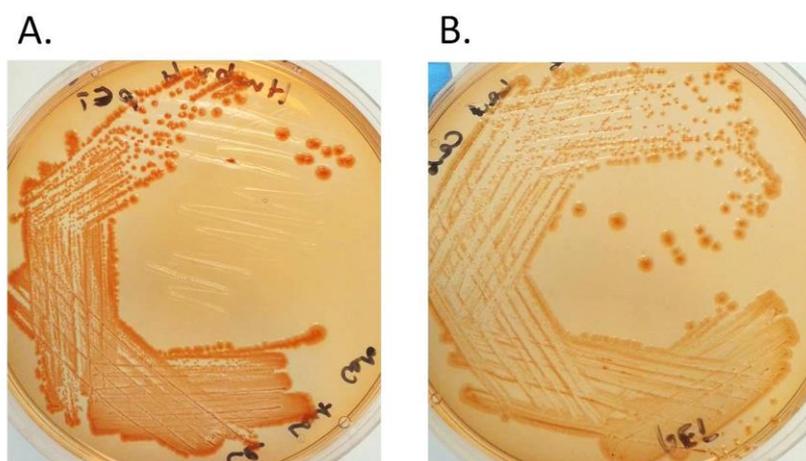
Figure 3.8 Congo red binding of *E. coli* cells expressing HmbR grown on Congo red inducing plates.

Cells were streaked out on BHI plates supplemented with lactose (0.05% w/v), ampicillin (100 µg/ml) and Congo red (0.005% w/v). **A.** BL21Star (DE3) cells with pET21-HmbR construct; **B.** BL21Star(DE3) with an empty pET21 vector.

Cells expressing *hmbR* gene reduced the concentration of haemin in the solution by up to 80%, whereas cells with an empty vector showed negligible binding. The highest level of haemin binding was observed up to 60 minutes post-induction and the ability of bacteria to bind haemin slowly decreased after extended induction. Previous observation of Congo red plates with bacteria

118

streaked after induction showed a drastically decreased number of colony forming units after induction indicating a toxic effect of *hmbR* expression on the cells. Moreover, there was a mixture of pink and red colonies indicating a loss of ability to express the gene by some bacteria. Those bacteria probably started to overgrow bacteria expressing the *hmbR* gene, hence the reduction of binding after extended induction. *E. coli* expressing *hmbR* did not show obvious binding of haemoglobin. A similar experiment performed by Simpson and co-workers (Simpson *et al.*, 2000) with overexpressed *hmuR* receptor from *Porphyromonas gingivalis* showed some level of haemoglobin binding (around 15%). However, the cell strain used for overexpression BL21(DE3)pLysE allowed a slower induction rate which could be more efficient for incorporation of the produced receptor in the membrane. This could result in more effective binding of haemoglobin to the cells.
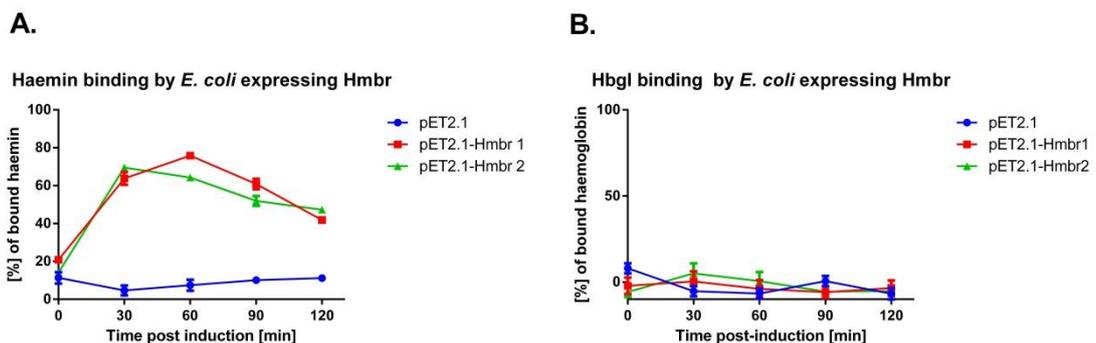
Figure 3.9 Binding of haemin and haemoglobin by *E. coli* expressing *hmbR*

BL21Star(DE3) cells with an empty pET21 vector and duplicate cultures expressing *hmbR* (pET21-Hmbr1 and pET21-Hmbr2) were taken at 30 minutes intervals post-induction and incubated with haemin or haemoglobin at final concentrations of 5 μM and 10 μM respectively.

3.2.3.2 HmbR interactions with Gly1ORF1

Gly1ORF1 interactions with recombinant HmbR were examined by a whole cell binding assay using *E. coli* BL21Star(DE3) producing HmbR as described in Materials and Methods Section 7.10.4. Western blot analysis of cell pellets and

supernatants after incubation with Gly1ORF1 showed reduction of Gly1ORF1 concentration in the supernatants incubated with cells expressing *hmbR*. No obvious reduction in the Gly1ORF1 concentration in the supernatant was observed for cells with an empty pET21 vector (Fig. 3.10 and 3.11 Panel A.). In the samples with cells harbouring pET21-HmbR construct there was a lower amount of Gly1ORF1 in the first 60 minutes post induction. After 90 minutes post-induction the amount of Gly1ORF1 in the supernatants was comparable to the samples with an empty vector. The statistical analysis shows that at 60 minutes post induction, the difference between samples with an empty pET21 and pET21-HmbR vector in depletion of Gly1ORF1 from the supernatants is statistically relevant (Fig. 3.11). The highest depletion of Gly1ORF1 in the supernatant observed 60 minutes post-induction corresponds with the results for haemin binding (Fig. 3.9, Panel A). This is probably an optimal induction time for the *hmbR* expression as extended inductions leads to the selection of cells that lose the ability to express *hmbr* gene, which was mentioned earlier in this chapter.
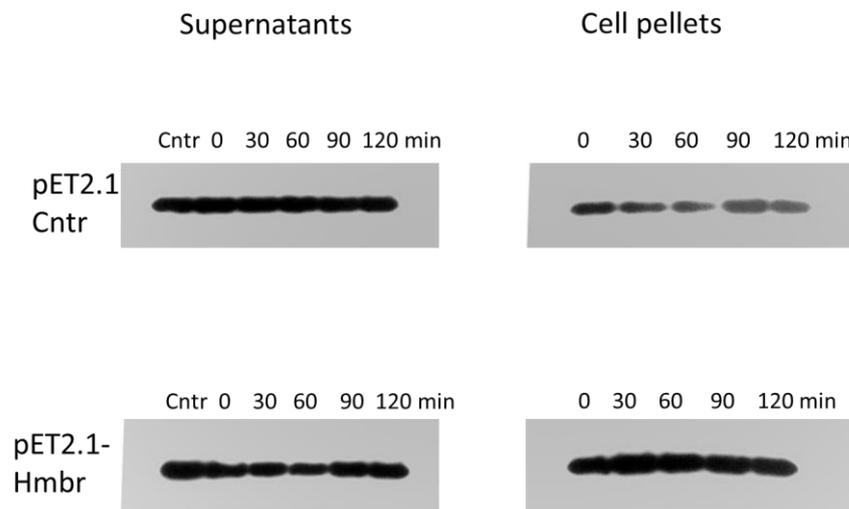
Figure 3.10 Binding of Gly1ORF1 by *E. coli* expressing *hmbR*.

E. coli cells with an empty pET21 vector and the pET21-HmbR construct were grown to a mid-log phase and induced with 0.1 mM IPTG. Samples at a final volume of 1 ml were incubated with Gly1ORF1 for the times indicated at the final concentration of 2.5 µM for 1 hour at 37°C. After incubation cells were centrifuged down and supernatants were removed and cell pellets were resuspended in 100 µl of TE buffer containing 1% (w/v) SDS and boiled for 5 min. Supernatants and cell lysates were analysed by western blot with the anti-Gly1ORF1 rabbit antibody. This figure is representation of three independent experiments.

The analysis of the cell lysates after incubation of Gly1ORF1 shows higher binding of Gly1ORF1 to the cells expressing *hmbR* at 30 minutes post induction. The differences can be explained by the fact, that analysis of the western blot images is only a semi-quantitative method and is affected by the quality of the transfer and detection with the antibodies. However, the data clearly show a difference in the binding between the cells expressing *hmbR* and cells with an empty vector (Fig. 3.11).

A.
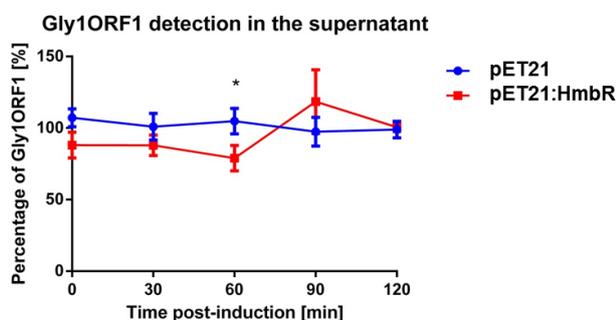
B.
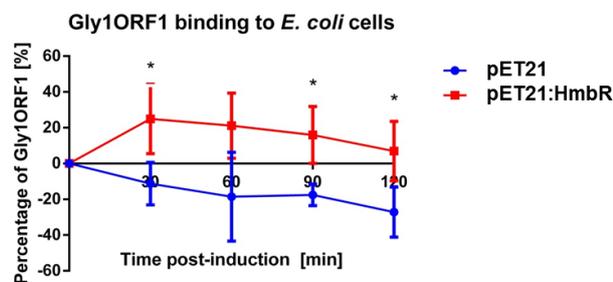
Figure 3.11 Gly1ORF1 binding by the *E. coli* cells expressing *hmbR*.

A semi-quantitative analysis of western blot images with ImageJ software. **A.** Analysis of depletion of Gly1ORF1 in the supernatants after incubation with induced *E. coli* cells containing an empty pET21 vector and pET21-HmbR construct. The percentage of Gly1ORF1 in the supernatant was calculated using a negative control (Gly1ORF1 solution with no cells added) as a 100% reference. **B.** Analysis of Gly1ORF1 binding to *E. coli* cells at the time points after induction. Samples taken before induction were used as a 0% reference. The graph shows results of 3 independent experiments with error bars showing SEM. The results with an asterix were statistically significant when analysed using the student t-test.

In summary, in this subsection *hmbR* was successfully cloned and expressed in *E. coli* BL21Star(DE3) cells and the haem binding and Congo red binding assays show that the receptor is functional. The data of the experiments with Gly1ORF1 binding to the whole *E. coli* cells expressing *hmbR* shows some level of Gly1ORF1 binding compared to the cells with an empty vector. This opens a

possibility of further studies on Gly1ORF1 being involved in the iron acquisition in *Neisseria* spp.

CHAPTER 4   IMMUNOLOGICAL STUDIES ON GLY1ORF1

4.3 Introduction

After many years of research, in January 2013 a new MenB vaccine – 4CMenB was licensed by the European Commission for children older than 2 months. The vaccine is a result of a novel approach in vaccine design, called reverse vaccinology. The vaccine candidates were selected after *in silico* search for surface exposed and potentially immunogenic antigens with a high degree of sequence similarity within the genome of the MC58 strain. The aim was to find the surface exposed antigens that would confer broad strain coverage.

According to the study by Vogel and co-workers, the 4CMenB vaccine could protect only against 78% percent of MenB strains in Europe (Vogel *et al.*, 2013). The data show clearly that there is still a need for the development of a more effective vaccine, with higher strain coverage and ideally, with cross-protective activity also to other serotypes that would enable eliciting herd-immunity similarly to widely used MenC vaccine. The search for surface-exposed antigens that are highly conserved in *Neisseria* spp*.* and are able to induce a long-term antibody response still continues. There are many genes in the neisserial genome that could meet the criteria but remain largely uncharacterised.

One such protein is Gly1ORF1 which is a subject of this study. Gly1ORF1 is a small, positively charged protein. According to the experiments performed in our lab, the protein expressed in *E. coli* is secreted in the culture supernatant and present in outer membrane vesicles (Meadows, 2004; Sathyamurthy, 2011). Arvidson *et al* reported that the protein was also present in outer membrane preparations from *N. gonorrhoeae* (Arvidson *et al.*, 1999). The analysis of neisserial isolates for the presence of Gly1ORF1 genes performed by Meadows showed that 96.7% of the tested isolates had the *gly1*ORF1 gene. The sequencing results revealed a high level of sequence conservation with many strains showing 100% sequence identity (Meadows, 2004).

In addition, an analysis of the gene prevalence in the neisserial genome database confirmed presence of the gene in sequences of over 200 clinical isolates with a high degree of sequence homology (Sayers, unpublished data).

The initial experiments done by Sathyamurthy showed that antibodies raised against Gly1ORF1 have detectable bactericidal activity against the MC58 strain *of N. meningitidis* (Sathyamurthy, 2011). The results encouraged further studies on Gly1ORF1 immunogenic properties.

Results from our laboratory suggested that Gly1ORF1 is immunogenic as the generation of anti-Gly1ORF1 sera resulted in a high antigenic response in rabbit. This chapter focuses on testing if high antibody levels against Gly1ORF1 could be detected in sera of patients diagnosed with meningococcal infection and if the incidence of the invasive disease resulted in the increase in anti-Gly1ORF1 antibody levels after a month long convalescence. Also, the study tried to establish if a carriage state could result in increased antibody levels to Gly1ORF1. In order to test that, sera of 96 patients with no suspected meningococcal infection were screened for an increased response to Gly1ORF1.

To test if antibodies raised against-Gly1ORF1 can confer a protective bactericidal activity, a serum bactericidal assay was performed with Gly1ORF1 anti-sera and affinity purified antibodies. This method is a gold standard for evaluation of potential vaccine candidates (Borrow *et al.*, 2005b). The principle of the method is to test if antibodies against the target antigen can kill bacteria in presence of the exogenous complement source.

In order to activate complement mediated killing via the classical pathway, an antigen must be surface exposed to allow the recognition by the protective antibodies. To evaluate if Gly1ORF1 meets this criterium, several approaches were undertaken to detect expression of the protein in several neisserial isolates including western blot analysis of outer membrane vesicles and whole cell lysates, Whole Cell ELISA and FACS analysis.

4.4 Results and Discussion

4.4.1 Levels of anti-Gly1ORF1 antibodies in sera of septic patients and volunteers with no suspected meningococcal infection

Sera from 6 patients (aged 22-43 years) with acute meningococcal septicaemia taken on admission to hospital and one month after infection were screened by enzyme-linked immune-absorbent assay (ELISA) for the presence of antibodies raised against Gly1ORF1 and NM MC58 outer membrane vesicles (OMVs) which served as a positive control for the assay (see Materials and Methods Section 7.11.2). The aim of the experiment was to establish whether exposure to meningococcal infection is associated with increased levels of antibodies to Gly1ORF1 during and after the infection.

Serum of one of 6 patients (Fig. 4.1; Patient 2) showed a response to Gly1ORF1 comparable to the response against OMVs. There was no obvious difference between samples taken during acute infection and one month post-infection. There was some increased response against Gly1ORF1 in Patient 4. Sera from those two patients were tested at higher dilutions to determine the antibody titres against Gly1ORF1 alongside a serum sample from a healthy volunteer with no history of meningococcal infection which was used as a control (Fig. 4.2).

Patient 2 showed a positive response to Gly1ORF1 and OMVs for serum dilutions up to 1:160. An antibody titre against Gly1ORF1 in patient 4 was determined to be 1:80 compared to the MC58 OMVs titre which was 1:320. The healthy volunteer exhibited a titre against OMVs higher than 1:500, whereas some response to Gly1ORF1 was observed only up to a 1:20 dilution. This finding was similar to the response of remaining septic patient samples tested in the initial screen. The presence of antibodies to Gly1ORF1 was detected in two out of 6 septic patients indicating that Gly1ORF1 is able to induce immunogenic response in some cases. No information was available about the strains causing the meningococcal infection in tested patients. It is possible that levels of expression of Gly1ORF1 vary from strain to strain and can be also influenced by the host-pathogen interactions, hence the low response to Gly1ORF1 was observed in the remaining patients.
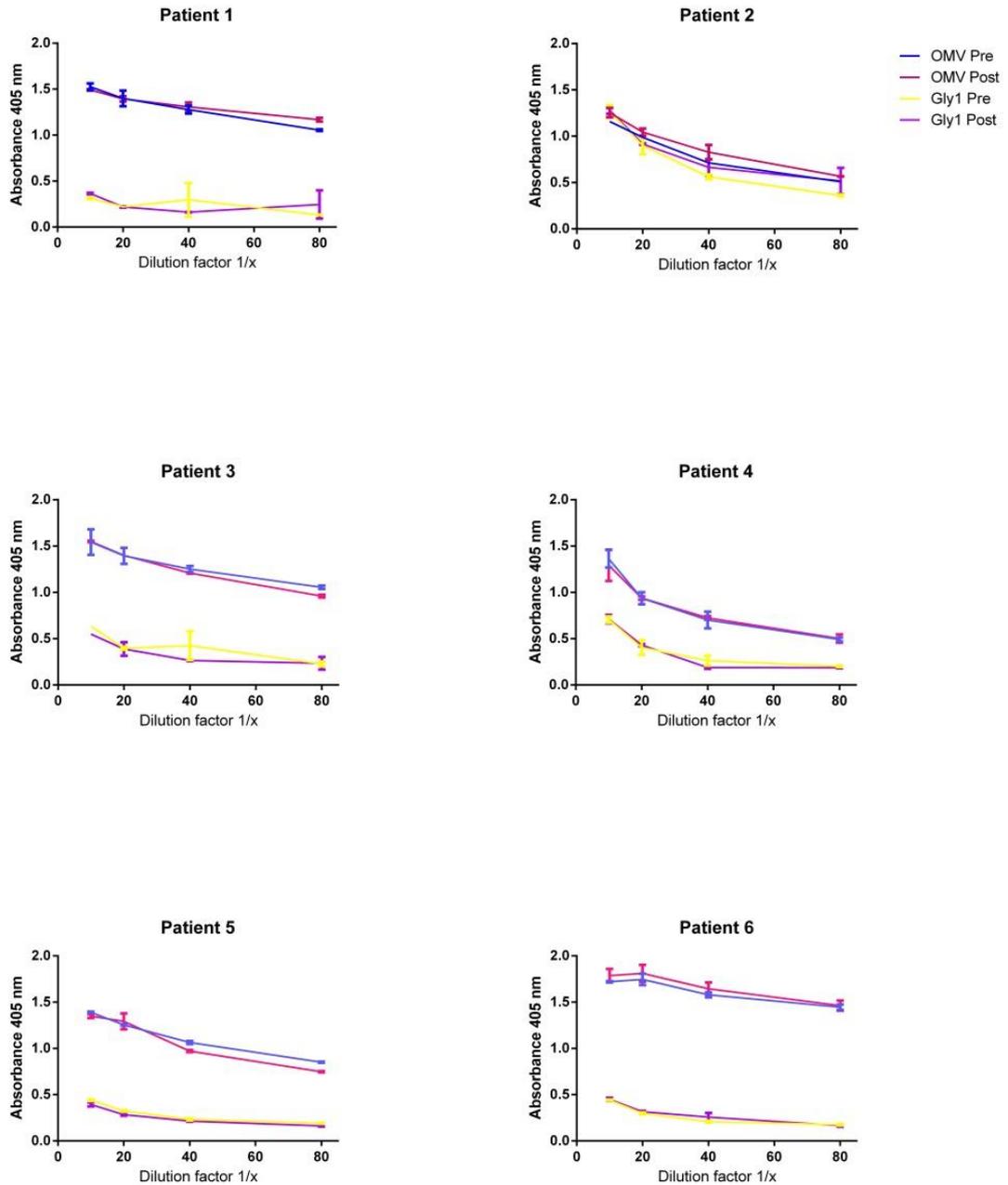
Figure 4.1 Results of initial ELISA screen for antibody response to Gly1ORF1in sera from 6 patients with acute septicaemia

Sera were taken on admission to hospital (PRE) and one month after infection (POST).Test performed in triplicate at each serum dilution. The anibody response to Gly1ORF1 (5 μg/ml) and NM MC58 outer membrane vesicles (OMV – 1 μg/ml) was measured in triplicate. The error bars indicate SEM.
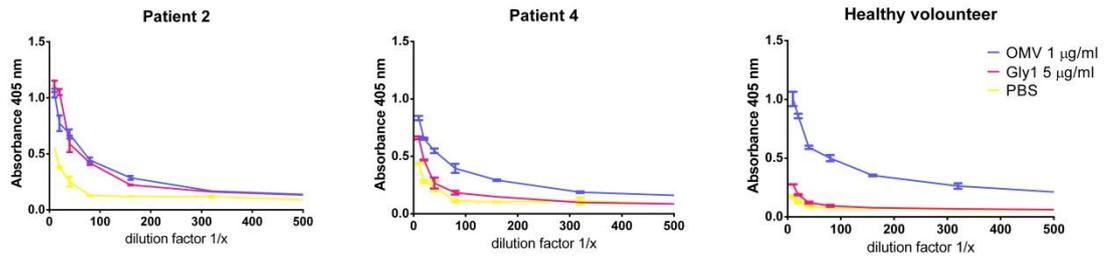
Figure 4.2 Results for ELISA antibody titres against Gly1ORF1 and MC58 OMV in human sera.

Sera from two septic patients and one healthy volunteer were tested in serial two-fold dilution in triplicate. Error bars indicate SEM.

In order to determine what is the prevalence of increased antibody response to Gly1ORF1 in the healthy population, sera of 96 patients with no suspected meningococcal infection were initially screened at 1:20 dilution to select samples with highest response for further analysis. The mean and standard deviation were calculated for the population of 96 serum samples. Samples with response to Gly1ORF1 higher than one standard deviation were tested at higher dilutions to determine antibody titres against Gly1ORF1. Outer membrane vesicles of NM MC58 were used as a positive control for the assay.

The tested samples were obtained from individuals at the age of 18-50. The report by Health Protection Agency shows that mean antibody titres against OMV from serogroup B meningococci vary between different age groups. The anti-OMV IgG reach a maximum level in individuals at the age of 19 (geometric mean titre over 70) and remain fairly high in older age groups (geometric mean titres varying from 50-75) (Trotter *et al.*, 2007a). The antibody titres vary from 45 to 75. It was expected that in the population of the tested samples there would be a high response to outer membrane vesicles.
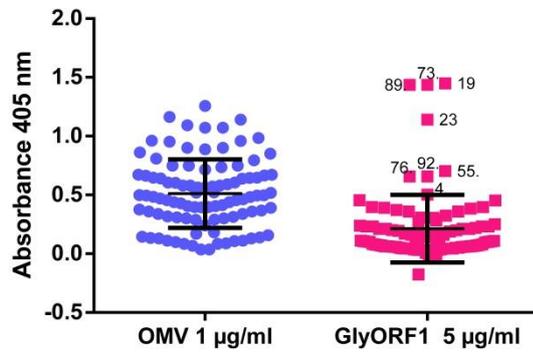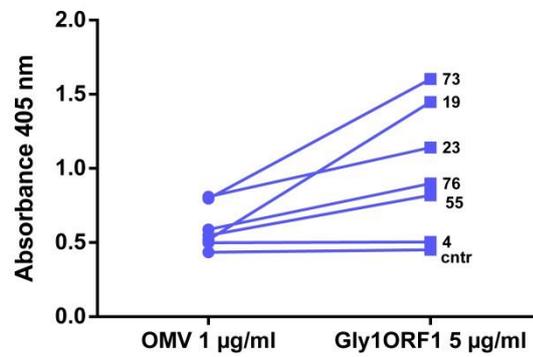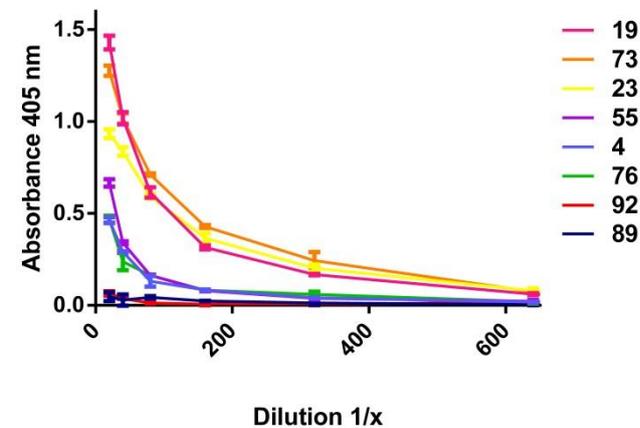
Figure 4.3 Analysis of presence of Gly1ORF1 antibodies in cohort of 96 patients with no suspected meningococcal infection.

A. An initial ELISA screen for samples with highest response to Gly1ORF1. MC58 OMV served as positive control. Sera were tested at 1:20 dilution. The graph shows means of three repeats and error bars show one standard deviation. **B.** Comparison of response to OMV and Gly1ORF1 in positive samples. **C.** ELISA: anti-Gly1ORF1 titre measurement in selected samples. The graph shows the means of three repeats with error bars showing SEM

128

In the initial screen, 7 out of 96 non-septic patients (7.3%) showed higher than one standard deviation response to Gly1ORF1. All selected samples had higher response to Gly1ORF1 than to control MC58 OMV apart from sample number 4 which was just on the border line of standard deviation (Fig. 4.3, Panel B.). Additional ELISA confirmed high anti-Gly1ORF1 antibody titres for 5 out of 7 of the selected samples, where 3 of them showed a response up to a serum dilution of 1:320 and 2 of them at a dilution of 1:160 (Fig. 4.3, Panel C.). It is possible that two of the samples which in the initial screen gave positive result were sensitive to freeze-thaw cycle and lost the ability to interact with Gly1ORF1.

In summary, 5 out of 96 samples (5.21%) showed higher than one standard deviation response to Gly1ORF1. Results from this screen showed that Gly1ORF1 antibody levels can be increased even in people who were never diagnosed with meningococcal infection. It is possible that positive samples came from individuals with previous episodes of *N. meningitidis* carriage. This indicates that exposure to Gly1ORF1 expressed in *N. meningitidis* can induce human antibody response.

4.4.2 Evaluation of serum bactericidal activity of anti-Gly1ORF1 antibodies

One of the essential properties of a vaccine candidate is its ability to induce production of protective antibodies against the target organisms. This can be assessed by a serum bactericidal assay (SBA). This gold standard method tests the ability of antibodies raised against a target molecule to induce complement mediated killing of the pathogenic bacteria.

Initial experiments in our lab showed that rabbit anti-Gly1ORF1 affinity-purified antibodies are able to induce serum bactericidal activity (Sathyamurthy, 2011). The aim of this study was to determine the SBA titres for the Gly1ORF1 antisera and test whether there is a difference in the antibody titres in control sera from mice injected with PBS. For the purpose of this experiment 6 mice were immunised with Gly1ORF1. Three control mice were immunised with PBS

(see Materials and Methods Section 7.11.1). Purification of the antigen for immunisation and initial screen of mouse anti-sera were performed by Dr Rachel Walton. Immunisation and mouse blood collection was performed at BioServe UK Ltd at the University of Sheffield.

The antibody titres of pooled sera against Gly1ORF1 from 6 immunised mice were measured by ELISA as described in Materials and Methods Section 7.11.1. Results showed that mouse Gly1ORF1 anti-sera a showed detectable response to the protein even above 10^4 serum dilution (Fig. 4.4)
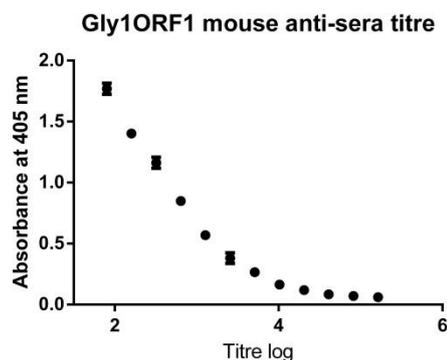
Figure 4.4 Antibody titres in mouse sera raised against Gly1ORF1.

Titres were determined by ELISA using purified Gly1ORF1 at the concentration of 5 µg/ml. An average of n=2 was plotted on the graph with error bars indicating SEM.

According to the Health Protection Agency standards, anti-sera tested with human complement that exhibit higher than 50% killing rate at 1:4 dilution or higher are considered bactericidal (Borrow *et al.*, 2005a). In order to test if sera raised against Gly1ORF1 were able to mediate complement mediated killing, a serum bactericidal assay (SBA) was performed as described in Materials and Methods Section 7.11.3. The SBA results with mouse anti-sera confirmed that Gly1ORF1 antibodies are able to induce complement-mediated killing of NM MC58 (Fig. 4.5; Tab. 4.1). De-complemented mouse anti-Gly1ORF1 sera showed killing above 50% for serum dilution up to 1:16 when tested with human complement. There was statistically significant difference in killing rate between anti-Gly1ORF1 and anti-PBS sera for dilutions up to 1:128. All controls with bacteria only, bacteria and human complement and anti-sera with heat-inactivated human complement were done. The results clearly show that anti-Gly1ORF1 antibodies have bactericidal effect on *N. meningitidis*.

SBA of mouse Gly1ORF1 anti-sera against WT NM MC58
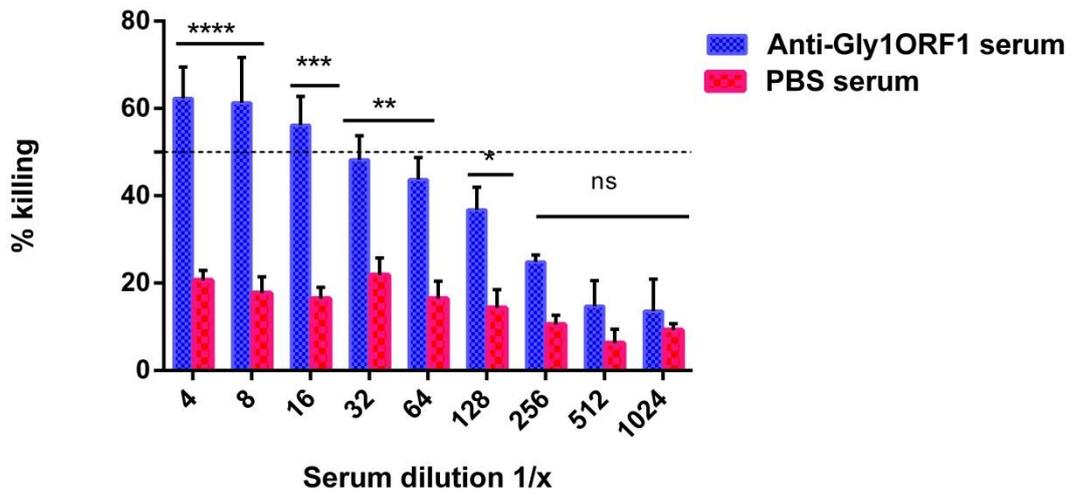
Figure 4.5 Summary of SBA results for mouse anti-Gly1ORF1 and PBS sera against wild-type NM MC58 tested with human complement.

Mean results plotted from 3 to 6 replicates for each dilution with error bars showing SEM. Statistical analysis of the results is shown in Table 4.2.

Table 4.1 Statistical data analysis of SBA results for mouse sera against Gly1ORF1 and PBS sera*

Sidak's multiple comparisons test	Mean Diff.	95% CI of diff.	Significant?	Summary	Adjusted P Value
Mouse anti-Gly1 serum – Mouse PBS serum					
4	41.65	17.24 to 66.06	Yes	****	< 0.0001
8	43.59	19.18 to 67.99	Yes	****	< 0.0001
16	39.74	15.33 to 64.14	Yes	***	0.0002
32	26.40	7.104 to 45.69	Yes	**	0.0022
64	27.29	7.992 to 46.58	Yes	**	0.0015
128	22.48	3.181 to 41.77	Yes	*	0.0132
256	14.36	-10.05 to 38.76	No	ns	0.5929
512	8.488	-15.92 to 32.89	No	ns	0.9685
1024	4.350	-20.06 to 28.76	No	ns	0.9998

* Data analysed using Two Way Anova test with Sidak's multiple comparison.

Another aim of the experiment was to check if there were differences between the wild-type and ΔGly1ORF1 isogenic mutant of NM MC58 in susceptibility to bactericidal activity of Gly1ORF1 antibodies. The limited volume of mice anti-sera available was not sufficient for further experiments, so rabbit anti-sera against Gly1ORF1 were tested in the SBA against the wild-type MC58 to provide further data.
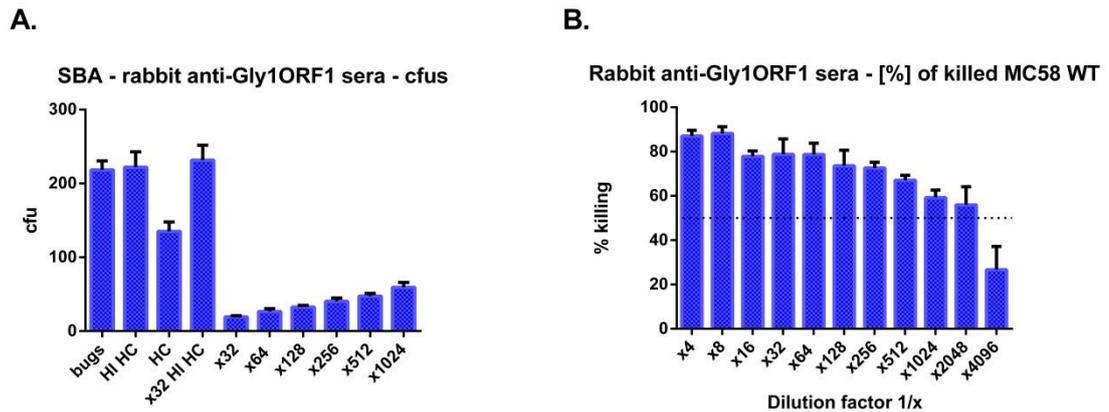
Figure 4.6 Serum bactericidal assay with rabbit Gly1ORF1 anti-sera.

A. Colony forming counts (CFUs) from one representative experiment showing controls with bacteria only (bugs), bacteria and heat inactivated human complement (HI HC), bacteria with human complement (HC) and bacteria with first tested dilution of rabbit antisera and heat inactivated human complement (x32 HI HC). **B.** Summary of experiments showing percentage of killed bacteria at serial two-fold dilutions of rabbit Gly1ORF1 anti-sera. The percentage of killing was calculated using the sample of bacteria with complement only (HC) as a baseline. Each dilution was tested in three to six replicates.

Rabbit Gly1ORF1 anti-serum showed very high SBA titres – up to 1:2048 (Fig. 4.6). The difference in SBA titres of mice and rabbit antisera against the same antigen could be caused by the differences in the solubility of the antigen used for immunisation. Rabbit anti-sera were raised against an insoluble form of Gly1ORF1 which can be more immunogenic (Hermeling *et al.*, 2004). In contrast, mice were injected with soluble Gly1ORF1 which could confer lower immunogenicity. According to Dr Myron Christodoulides from the University of Southampton (verbal communication) rabbit anti-sera can also show high background in serum bactericidal assays. Rabbit anti-Gly1ORF1 sera were

generated before the experiment was planned and no sample of serum before immunisation was available for testing to compare SBA before and after immunisation. Thus, it was impossible to check if unimmunised rabbit's serum possessed bactericidal activity that could add to Gly1ORF1 antibodies bactericidal effect on *Neisseria meningitidis.*

To compare wild-type and Gly1ORF1 knock-out MC58 susceptibility to bactericidal effect of Gly1ORF1, rabbit antibodies were affinity purified using Gly1ORF1 immobilised on cyanogen-bromide activated agarose beads as described in Materials and Methods Section 7.10.3.2). Purified antibodies were tested at two-fold serial dilutions starting from 1:20. Results for the assay are shown in Figure 4.7 and Table 4.2.

Figure 4.7 Serum bactericidal assay results: comparison of bactericidal activity of anti-Gly1ORF1 rabbit affinity purified antibodies against wild-type and ΔGly1ORF1 MC58 mutant.

Results show the means from three individual experiments with error bars showing SEM. Results of the statistical analysis of the results is shown in Table 4.2.

Table 4.2 Statistical analysis of SBA results for rabbit polyclonal antibodies tested against WT and Gly1ORF knock-out*.

Sidak's multiple comparisons test	Mean Diff.	95% CI of diff.	Significant?	Summary	Adjusted P Value
WT - KO					
20	24.55	4.382 to 44.71	Yes	*	0.0119
40	20.56	0.3980 to 40.73	Yes	*	0.0441
80	31.58	11.41 to 51.74	Yes	**	0.0011
160	32.31	12.15 to 52.48	Yes	***	0.0009
320	28.86	4.169 to 53.56	Yes	*	0.0165
640	30.40	5.705 to 55.10	Yes	*	0.0109

*Data analysed using Two Way Anova test with Sidak's multiple comparison.

Results of SBA assay showed that the difference in killing between wild-type and Gly1ORF1 knock-out MC58 was statistically significant. The wild-type was more susceptible to killing than the knock-out MC58. There was some background killing of the knock-out in the first tested dilutions. However, as the antibodies used for the assay were at high concentration (0.8 mg/ml) they could show some hydrophobic interaction with the surface of knock-out bacteria and activate complement-mediated killing in a non-specific manner. At 1:80 dilution, the antibodies showed no killing of the knock-out and at higher dilutions bacteria were not affected by the presence of the antibodies and their number increased after incubation, whereas wild-type MC58 numbers at these dilutions were still reduced. This was presumably because the knock-out bacteria were able to continue to replicate in the course of the assay whereas wild type had been killed. Although percentage of killed wild-type NM MC58 is significantly higher than killing of Gly1ORF1 knock-out, it did not reach the 50% killing rate and is lower than the results previously described by Sathyamurthy. It is possible that affinity purification of antibodies in harsh conditions resulted in some level of denaturation which can vary from batch to batch. Purified antibodies were still detecting Gly1ORF1; however, it could interact less efficiently with the complement system and be less effective in inducing killing of MC58. Also, the main drawback of the SBA method is low reproducibility due to the differences in the antibody levels between complement donors, media used for bacterial growth before the assay and length of the incubation. (Andrews and Pollard, 2014). The main outcome of this experiment is confirmation of the ability of

Gly1ORF1 antibodies to induce complement-mediated killing by targeting Gly1ORF1. It can be also concluded that Gly1ORF1 is expressed on the surface of the cells as the wild-type MC58 was interacting with the anti-Gly1ORF1 antibodies much more effectively than the Gly1ORF1 knock-out.

4.4.3 Confirmation of surface localisation of Gly1ORF1 in *N. meningitidis*

To effectively induce complement mediated killing, an antigen used for immunisation must presumably be surface exposed on the pathogen to enable activation of the complement cascade. In order to check if Gly1ORF1 can be detected on the surface of the bacteria, outer membrane vesicles were purified and subjected to western blot with anti-Gly1ORF1 antibodies as well as whole-cell lysates. Another approach included a whole-cell ELISA with immobilised pasteurised neisserial cells and FACS analysis of neisserial cells with Gly1ORF1 antibodies and fluorescently labelled secondary antibody.
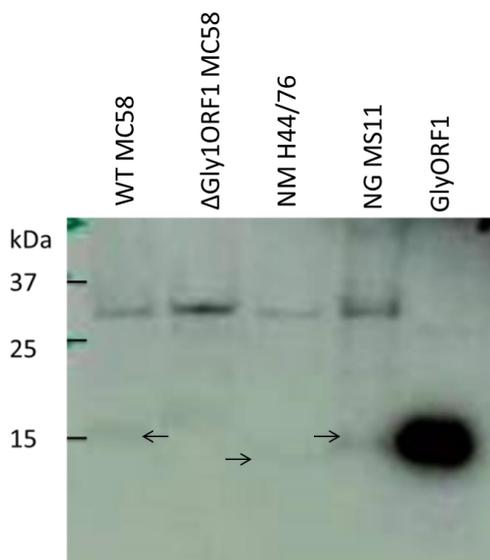
Figure 4.8 Detection of Gly1ORF1 by western blot in outer membrane vesicles of neisserial isolates.

The OMVs were isolated from: wild type NM MC58, isogenic Gly1ORF1 mutant (ΔGly1ORF1 MC58), NM strain H44/76 and *N. gonorrhoeae* MS11. Purified Gly1ORF1 C-histidine tag was used as a reference. A 10 μg aliquot of purified OMV was loaded in each lane. Arrows indicate protein bands at the similar size to Gly1ORF1

Analysis of western blots with outer membrane vesicles isolated from different strains of *Neisseria* spp. showed very faint bands corresponding to Gly1ORF1 size for WT MC58 and NG MS11 (Fig. 4.8). Gly1ORF1 knock-out outer membrane vesicle showed proteins at higher molecular weight than Gly1ORF1, but no protein at the same size. Outer membrane vesicles of strain H44/76 showed a band running slightly lower than proteins detected in WT MC58 and

135

NG MS11. Higher molecular weight proteins at approximately 30 kDa were observed for all tested strains probably showing some nonspecific interactions. The results indicate that Gly1ORF1 is probably expressed at a very low level and is not readily detectable by western blot or the antibodies are not sensitive enough to detect the antigen under the conditions used to culture the bacteria.
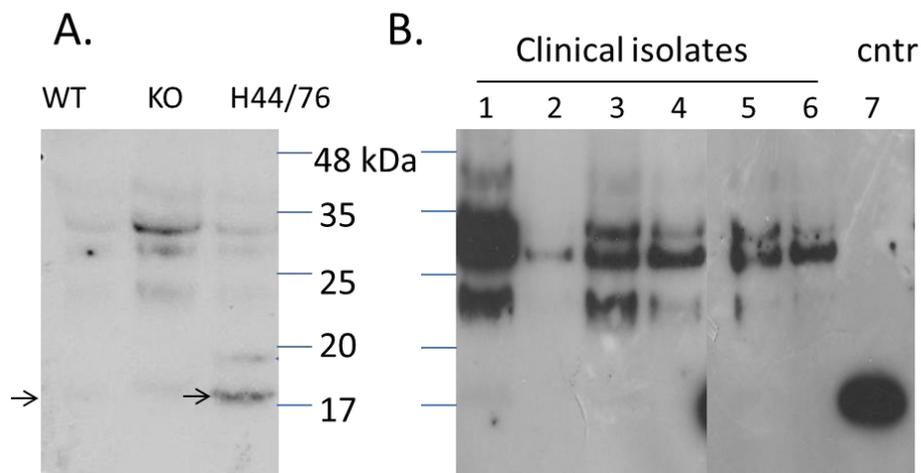
Figure 4.9 Immunoblot of the whole cell lysates of *N. meningitidis* strains with anti-Gly1ORF1 antibodies.

Panel A. contains lysates of laboratory strains: wild type NM MC58 (Lane WT), isogenic mutant ΔGly1ORF1 MC58 (Lane KO) and NM strain H44/76. In Panel B, **Lanes 1-6** represent cell lysates of clinical isolates NM1, 2, 3, 5, 6 and 7 respectively; **Lane 7** – shows control with purified Gly1ORF1. Protein bands at Gly1ORF1 size are indicated with arrows.

Western blot analysis of whole-cell lysates showed a detectable band at approximately 17 kDa that was not detected in ΔGly1ORF1 MC58 (Fig. 4.9, Panel A, Lane WT and H44/76). It could be unprocessed version of Gly1ORF1. Analysis of ten times concentrated liquid culture supernatants did not show obvious presence of Gly1ORF1 (data not shown). It was previously described that in NG MS11 Gly1ORF1 was only detected in concentrated outer membrane vesicles when 20 µg of isolated outer membrane vesicles was loaded in a lane (Arvidson *et al.*, 1999). However, due to the high viscosity of outer membrane fraction the amount of OMV was limited to 10 µg in this experiment. According to the literature, Gly1ORF1 expression could be up-regulated upon contact with host cells (Deghmane *et al.*, 2003) so one can conclude that the growth

conditions of analysed neisserial strains were not optimal for high levels of Gly1ORF1 expression.

A whole-cell ELISA experiment with immobilised cell suspensions of neisserial isolates (see Materials and Methods Section 7.11.4) gave very low signal with anti-Gly1ORF1 antibodies (Fig. 4.10). The capsule knock-out version – C13 gave slightly higher signal than MC58 WT indicating that presence of the capsule may mask Gly1ORF1 interactions with antibodies. The whole-cell ELISA experiment did not give convincing results for the detection of Gly1ORF1.
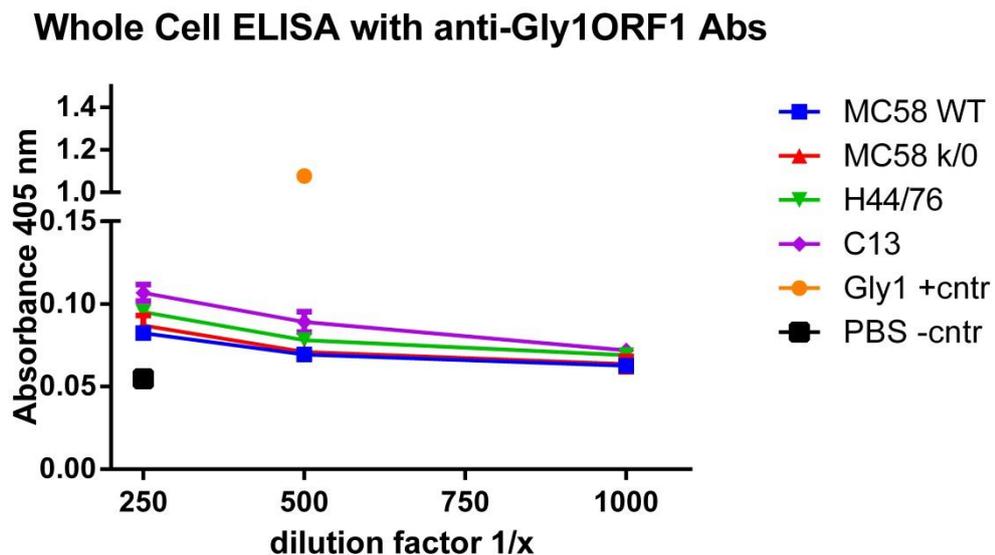
Figure 4.10 Detection of Gly1ORF1 on the cell surface by a whole-cell ELISA approach.

Pasteurised cell suspensions of NM MC58, Gly1ORF1 knock-out MC58, NM H44/76, capsule deficient mutant-C13 and purified Gly1ORF1 (5 µg/ml) as a positive control for the assay were tested with different dilutions of affinity purified anti-Gly1ORF1 antibody stock at the concentration of 0.8 mg/ml.

4.4.4 FACS analysis of neisserial isolates

Another approach to detect Gly1ORF1 on the surface of the cells was undertaken using the method described by Hung and co-workers (Hung *et al.*, 2013). Neisserial strains were first treated with ethanol to permeabilise the capsule and probed with antibodies against the target antigen followed by addition of fluorescently labelled secondary antibody as described in Materials and Methods Section 7.11.5). In this experiment purified rabbit anti-Gly1ORF1 antibodies were used. As a positive control for the experiment antibodies against alpha protein of IgA1 protease were used. Previous experiments in our lab performed by Parsons (Parsons, 2003) demonstrated that the majority of the tested neisserial strains showed considerable shifts in fluorescence when exposed to anti-alpha affinity purified rabbit antibodies.
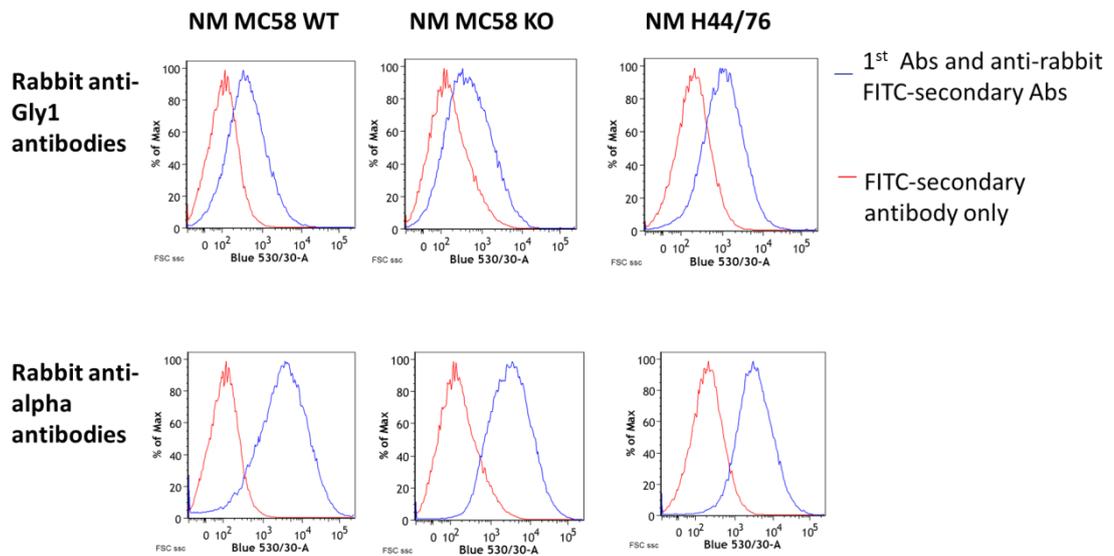
Figure 4.11 FACS analysis of presence of Gly1ORF1 on the cell surface.

MC58 WT, Gly1ORF1 knock-out and NM H44/76 were exposed to purified rabbit anti-Gly1ORF1 antibodies and FITC-labelled secondary antibody. Anti-alpha antibodies were used as a positive control of the assay.
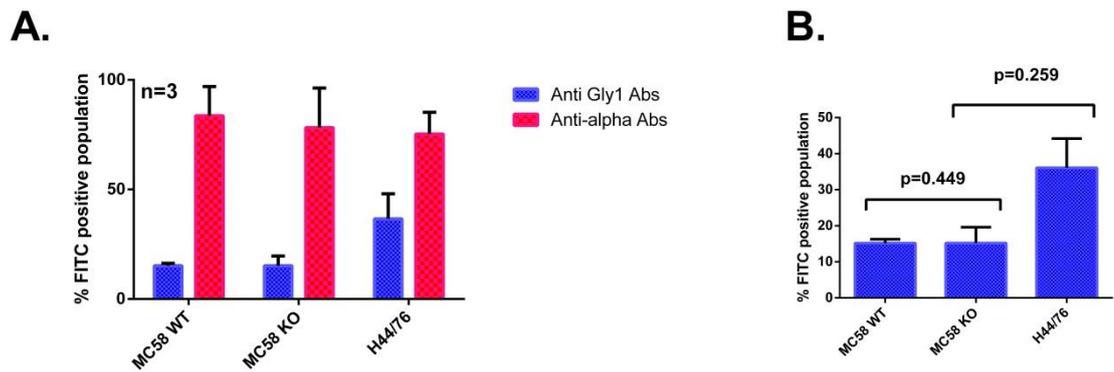
Figure 4.12 Analysis of FACS results with three strains of *N. meningitidis*.

A. Comparison of percentage of FITC-positive population after incubation with anti-Gly1ORF1 or anti-alpha antibodies (positive control) and FITC-labelled secondary antibody. **B.** FACS results with anti-Gly1ORF1 antibodies with results of statistical analysis using t-test. Data shows results of three independent experiments with error bars showing SEM.

Results of the experiment showed very low levels of fluorescence shift when testing neisserial strains with anti-Gly1ORF1 antibodies. Wild-type MC58 showed comparable level of the shift to Gly1ORF1 knock-out. Another tested strain – H44/76 showed higher level of fluorescent shift compared to wild-type and knock-out MC58, however the difference was not statistically significant. (Fig. 4.11 and 4.12). In contrast, anti-alpha antibodies showed almost complete fluorescence shift for all of the strains.

Six clinical strains were also tested using the FACS method to check if there is a difference in levels of detectable Gly1ORF1 in carriage strains and strains from systemic blood infection. Isolates NM1 to 5 were associated with carriage whereas isolates NM6 and NM7 were invasive strains isolated from blood. All clinical strains showed very low (lower to ΔGly1ORF1 MC58) shift in fluorescence when probed with anti-Gly1ORF1 antibodies and almost complete shift with anti-alpha antibodies (apart from isolate NM7 which show only partial shift). A negative control with *Streptococcus* isolate showed no shift with either of the tested antibodies (Fig. 4.13, Fig. 4.14).
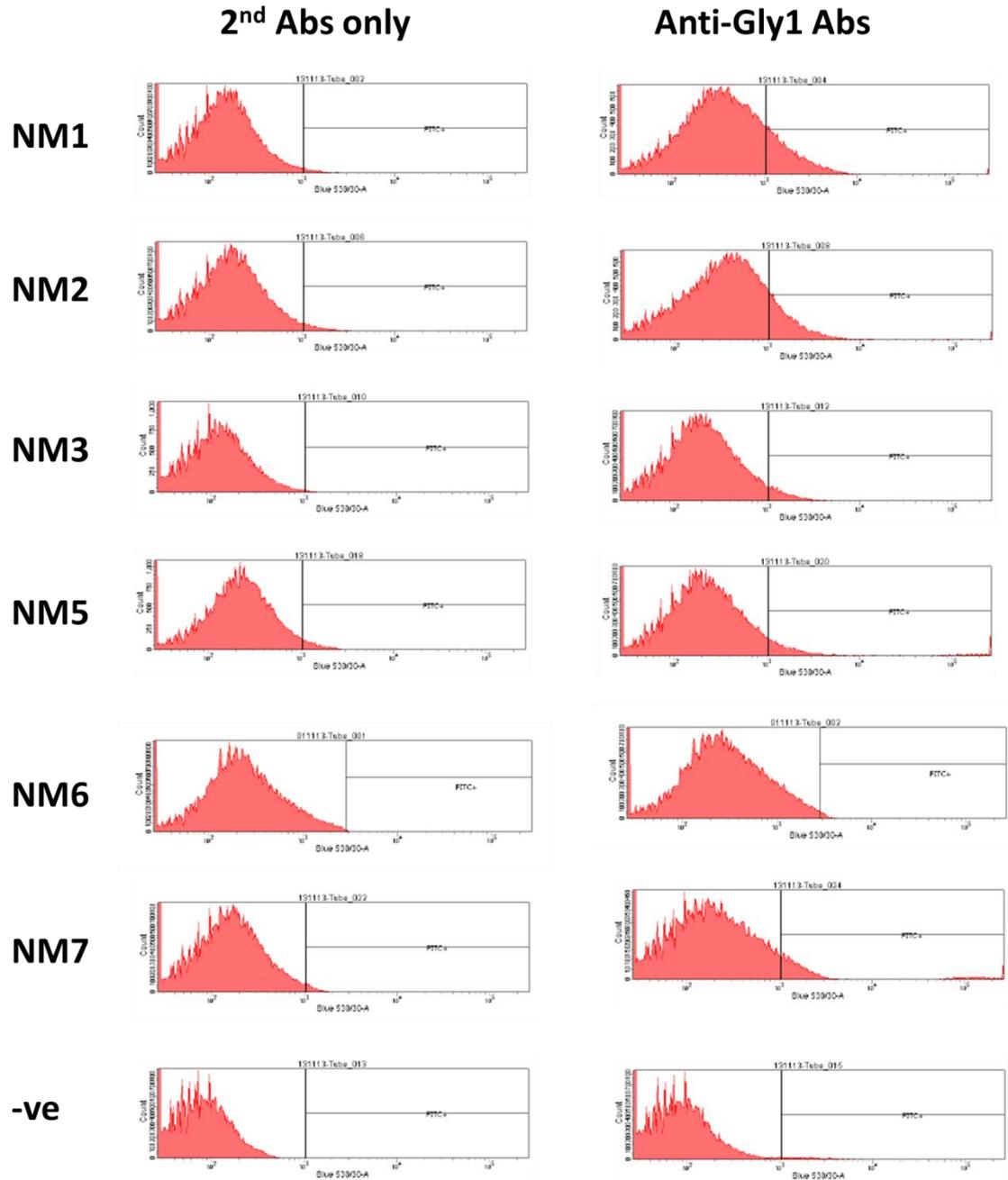
Figure 4.13 FACS analysis of Gly1ORF1 presence on the surface of neisserial clinical isolates.

The bacterial suspensions were first treated with ice-cold ethanol to permeabilise the capsule and then probed with anti-Gly1ORF1 antibody, followed with anti-rabbit FITC labelled secondary antibody. The data were collected on FACS LSRII machine at the Flow Cytometry Core Facility at the University of Sheffield. The negative control contained suspension of *Streptococcus* spp. treated in the same way as neisserial clinical isolates.
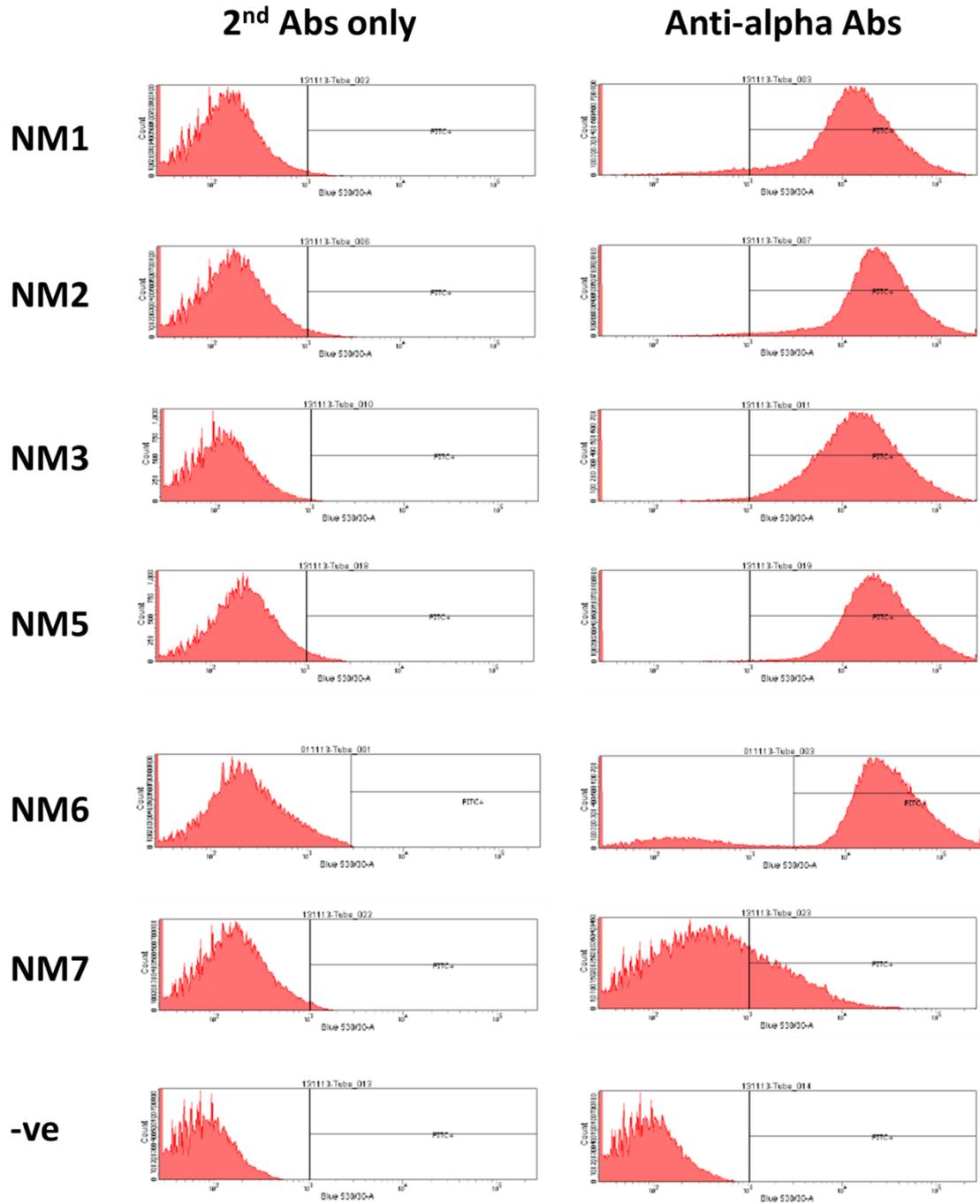
Figure 4.14 FACS analysis of alpha protein presence on the surface of neisserial clinical isolates.

The bacterial suspensions were also treated with ethanol first, then probed with anti-Gly1ORF1 antibody, followed with anti-rabbit FITC labelled secondary antibody. The data were collected on FACS LSRII machine at the Flow Cytometry Core Facility at the University of Sheffield. All isolates apart from NM7 showed complete shift in fluorescence. The shift was only partial, which might indicate lower levels of protein expression in this isolate. The negative control contained suspension of *Streptococcus* spp. treated in the same way as neisserial clinical isolates.

To examine whether *gly1ORF1* gene was present in the genomes of the clinical isolates, genomic DNA was used as a template for PCR with primers designed for *gly1ORF1* region: Gly1KOF- 5'-GCCGACGGCAAAACGGTTCA-3' and Gly1KOR-5'-CCAAACCGGCAGGCGCAAA-3'. Analysis of PCR products showed presence of the fragment at the same size as in MC58 WT for all clinical isolates and H44/76 strain. Gly1ORF1 knock-out, as expected, was at higher molecular weight due to the insertion of erythromycin cassette performed by Sathyamurthy (Sathyamurthy, 2011) (Fig. 4.15).
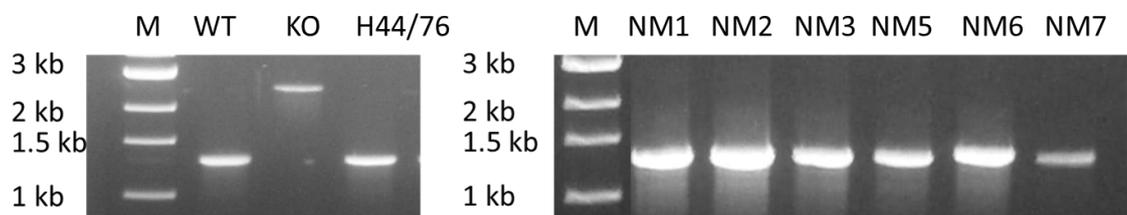
Figure 4.15 PCR products using genomic DNA of neisserial isolates and primers designed for Gly1ORF1 region.

Genomic DNA was isolated from the neisserial laboratory strains: wild type MC58 (WT), ΔGly1ORF1 MC58 (KO) and H44/76 and clinical isolates: NM1, NM2, NM3, NM5, NM6 and NM7. Lane M contains a DNA size marker.

4.4.5 SBA assay: MC58 WT vs. H44/76

A higher fluorescence shift of H44/76 with Gly1ORF1 antibodies compared to MC58 and western blot analysis of cell lysates and outer membrane fractions with detectable bands corresponding to Gly1ORF1 size indicates that the levels of *gly1ORF1* expression might be higher in this isolate than in MC58 and can reach detection level. To investigate whether higher levels of Gly1ORF1 protein in H44/76 strain correlates with higher susceptibility to complement mediated killing, a serum bactericidal assay with H44/76 and MC58 strain was done as described before.

Although there was a visible difference between MC58 and H44/76 in the percentage of killed bacteria at antibody dilutions up to 1:80, an analysis showed that the differences are not statistically significant (Fig. 4.16, Tab. 4.3).
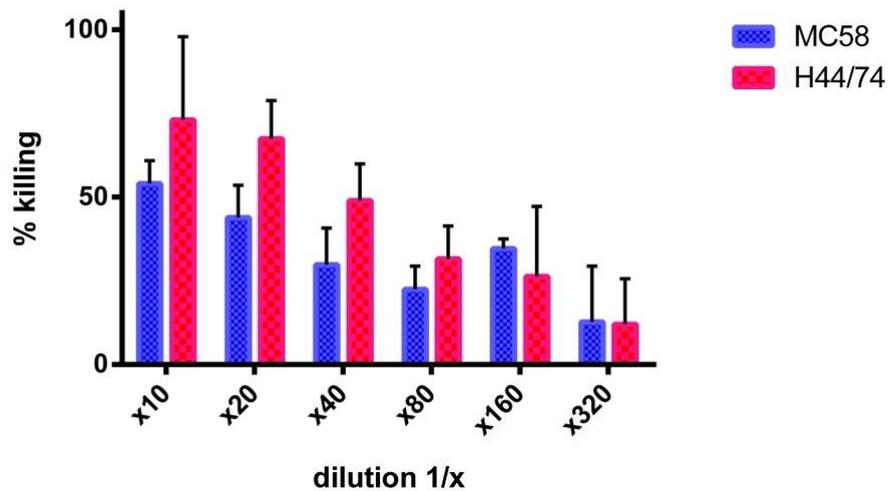
Figure 4.16 Serum bactericidal assay with anti-Gly1ORF1 antibodies – comparison of MC58 and H44/76 strains.

The data show results from three individual experiments. Anti-Gly1ORF1 antibodies stock at 0.8 mg/ml concentration was used. Statistical analysis using Sidak's multiple comparison test showed that the differences in percentage of killing for all antibody dilutions are not statistically significant. Error bars show SEM. The results of the statistical analysis are shown in Table 4.3.

Table 4.3 Statistical analysis of SBA test for strains MC58 and H44/76*

Sidak's multiple comparisons test	Mean Diff.	95% CI of diff.	Significant?	Summary	Adjusted P Value
MC58 - H44/74					
x10	-18.99	-54.85 to 16.87	No	ns	0.5932
x20	-23.53	-55.61 to 8.543	No	ns	0.2423
x40	-19.22	-51.30 to 12.86	No	ns	0.4565
x80	-9.132	-41.21 to 22.94	No	ns	0.9611
x160	8.327	-23.75 to 40.40	No	ns	0.9751
x320	0.6544	-38.63 to 39.94	No	ns	> 0.9999

Data analysed using Two Way Anova test with Sidak's multiple comparison.

In summary, Gly1ORF1 meets criteria for a vaccine candidate with regards to ability to induce antibody response in some individuals during meningococcal infection and there is some level of response to Gly1ORF1 in sera from population of patients with no meningococcal disease which could be a result of carriage. As serum bactericidal activity is regarded as a serological surrogate of protection against meningococcal infection (Borrow *et al.,* 2005b), (Frasch *et al.,*

2009), it might be indicative that antibodies against Gly1ORF1 might have protective effect. The methods used for detecting Gly1ORF1 on the bacterial surface gave inconclusive results probably due to the low sensitivity of affinity-purified antibodies and/or low levels of *gly1ORF1* expressed under the conditions used to culture the bacteria in the laboratory. The observation that WT MC58 is much more susceptible to complement killing upon addition of anti-Gly1ORF1 antibodies indicates that the antigen is expressed and surface exposed in MC58 WT. It can be also concluded that expression of Gly1ORF1 can be induced in the presence of human serum proteins during SBA assay or other host related factors. It would be worth exploring in which conditions the synthesis of Gly1ORF1 is up-regulated. An alternative approach for detection of *gly1ORF1* expression could be a real-time PCR, which in theory is more sensitive. Bacteria grown in the optimal conditions for Gly1ORF1 expression could be than tested again with methods described in this study. Another approach would be generation of monoclonal antibodies with higher affinity to Gly1ORF1 which would allow detection of much lower protein levels. The project to generate monoclonal antibodies was undertaken in our lab by Dr R. Walton with the help of BioServe UK Ltd, however, hybridoma clones generated either lost the ability to produce the antibodies or the antibodies were found to interact with the histidine tag and linker region of the recombinant protein used to immunize mice (data not shown).

4.4.6 Immunological studies on IgA1 alpha protein

Antibodies against IgA1 alpha protein were used in this study as a positive control for FACS analysis of Gly1ORF1 presence on the surface of the neisserial strains. FACS results with anti-alpha antibodies showed in all tested neisserial isolates almost complete shift of fluorescence indicating detectable levels of protein present on the surface of the cell (Fig. 4.20). These interesting results prompted further immunological studies on this protein with similar methods used for Gly1ORF1 analysis. Initial analysis included testing human sera for presence of anti-alpha antibodies, whole-cell ELISA, western blot analysis of cell lysates and outer membrane vesicles as well as ability to induce serum bactericidal activity.

ELISA with sera from two septic patients and one healthy volunteer showed high antibody levels to alpha protein (Fig. 4.17). Two septic patients had antibody titres comparable to a response to outer membrane vesicles. Surprisingly, the healthy volunteer showed a much higher response to alpha protein exceeding a serum dilution of 1:1280. This may indicate that antibody production against alpha protein can be induced during carriage.
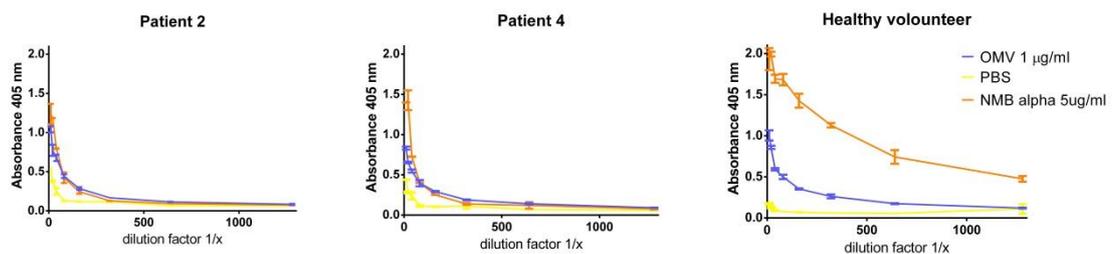
Figure 4.17 ELISA analysis of antibody response against purified NMB alpha protein (5 µg/ml) and OMVs (1 µg/ml)

Sera from two septic patients and one healthy volunteer were tested in serial two-fold dilutions.Error bars show SEM. Controls PBS are shown in the graph. Means of two replicates are plotted with error bars showing SEM, n=3.

It was already demonstrated that alpha protein can be detected with anti-alpha antibodies by FACS. The results were also confirmed by the whole-cell ELISA method. MC58 WT and H44/76 showed comparable levels of detected alpha protein. Higher response was observed for capsule deficient mutant C13. Similar effect was observed in the whole-cell ELISA with anti-Gly1ORF1 antibodies. Surprisingly, Gly1ORF1 knock-out mutant showed much higher response to alpha antibodies than other tested strains (Fig. 4.18). Gly1ORF1 knock-out was cultured on CBA plates containing erythromycin and it is possible that although the bacteria had an erythromycin resistance gene, the presence of the antibiotic affected bacterial gene expression profile or bacterial surface.
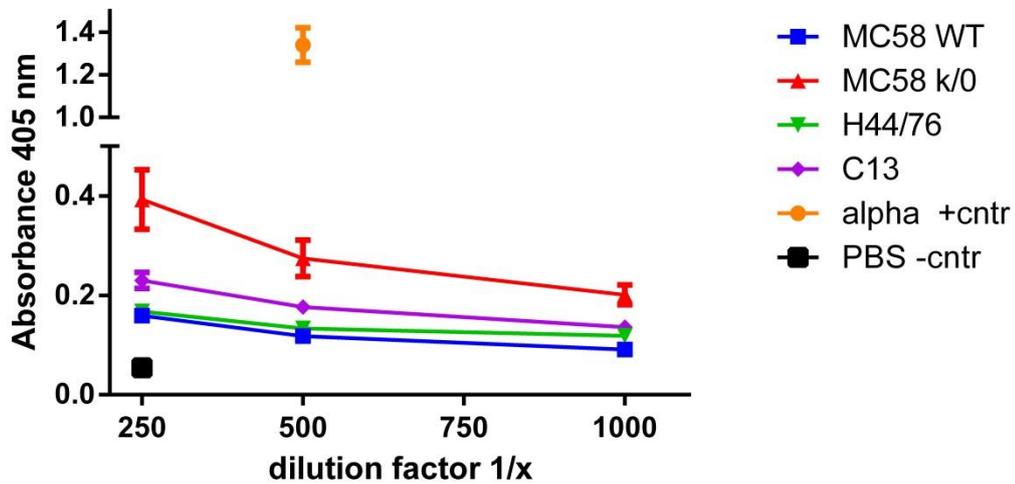
Figure 4.18 Detection of alpha protein on the cell surface by whole cell ELISA approach.

Pasteurised cell suspensions of NM MC58, Gly1ORF1 knock-out MC58, NM H44/76, capsule deficient mutant of C13 and positive control with purified immobilised alpha protein were tested with different dilutions of affinity purified anti-alpha antibodies at the stock concentration of 0.11 mg/ml.

Previously performed western blots with anti-alpha antibodies using cell lysates of the number of clinical isolates detected expression of alpha protein in majority of them (Parsons, 2003). Isolates from the laboratory stocks and clinical isolates analysis were similarly tested for presence of alpha protein. All isolates apart form H44/76 and NM7 showed detectable levels of a protein at the size of long alpha protein (Fig. 4.19).
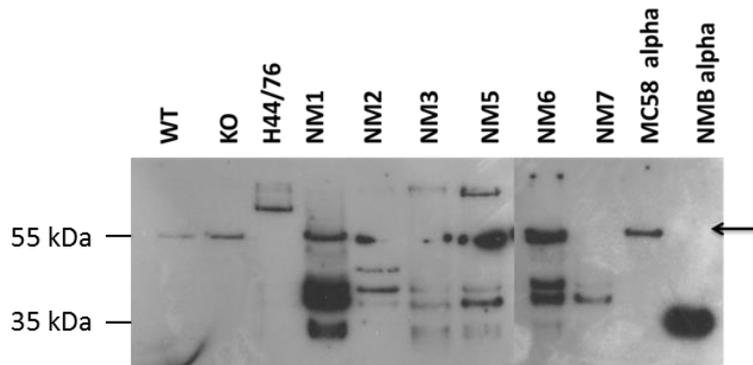
Figure 4.19 Immunoblot of cell lysates of neisserial isolates with anti-alpha antibodies.

Cell lysates of neisserial laboratory isolates: wild type MC58 (WT), ΔGly1ORF1 MC58 (KO), H44/76 and six clinical isolates (NM1-NM7) were separated in SDS-PAGE gel next to the reference lanes containing purified MC58 alpha protein and NMB alpha protein. Alpha proteins were detected by western blot using anti-alpha polyclonal antibodies. An arrow indicates the bands corresponding to the long alpha protein. Other proteins with lower molecular weight may be a product of auto-proteolytic activity of IgA1 protease resulting in shorter fragments of alpha protein. Proteins with higher molecular weight than long alpha protein may be a result of incomplete proteolytic cleavege such as γ-alpha fragments.

Finally, the serum bactericidal activity of anti-alpha antibodies was tested using affinity-purified rabbit antibodies. Results of the assay showed significant killing of NM MC58 WT with an antibody dilution at 1:20 (Fig.4.20). Further dilutions of the antibodies showed a reduction in percentage of killed bacteria. Controls with heat-inactivated human complement and first dilution of the antibody showed that the antibodies itself have no bactericidal activity as well as human complement itself as the colony forming units counts are similar for all the controls. It is worthy of note, that the percentage of killing is much higher than with anti-Gly1ORF1 antibodies even though the starting concentration of the antibodies was lower. Anti-Gly1ORF1 antibodies stock concentration was 0.8 mg/ml and anti-alpha antibodies only 0.11 mg/ml.
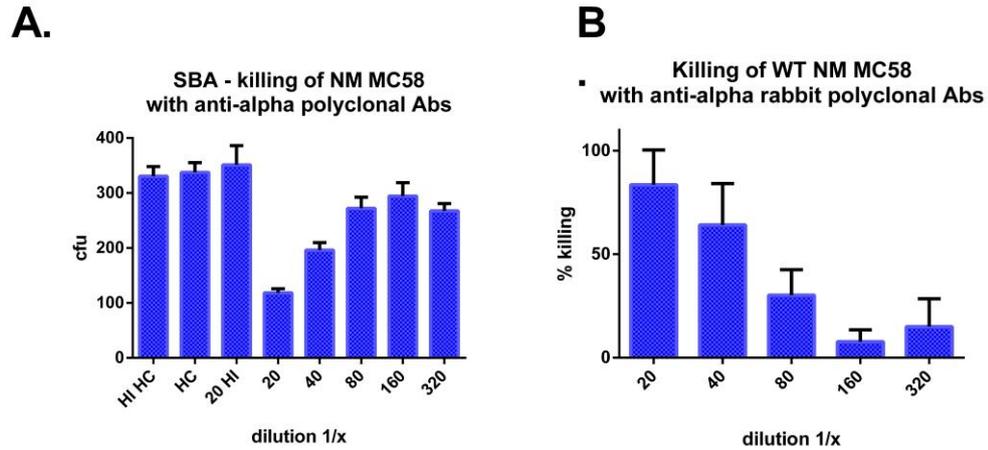
Figure 4.20 Serum bactericidal activity of rabbit anti-alpha affinity purified antibodies with human complement.

A stock concentration of anti-alpha antibodies used for the assay was 0.11 mg/ml. **A.** Changes in cfu counts upon treatment with anti-alpha antibody and human complement. Controls with heat-inactivated complement (HI HC), active human complement (HC) and heat-inactivated human complement with first dilution of the antibodies (20 HI HC) are shown. The data shown are from one out of three individual experiments. **B.** Mean percentage of killed bacteria from three individual experiments with affinity purified anti-alpha antibody. Error bars show SEM.

The higher efficacy of anti-alpha antibodies in the SBA can be explained by higher concentrations of alpha-protein on the cell surface. This was confirmed by ELISA, western blots and FACS analysis performed in this study. IgA1 protease was previously identified in microarray and proteomic analysis of neisserial proteins (Dietrich *et al.*, 2003; Echenique-Rivera *et al.*, 2011; Ferrari *et al.*, 2006) and analysis of the protein content of outer membrane vesicles (Lappann *et al.*, 2013), whereas Gly1ORF1 was not mentioned in those studies probably due to the low level of expression.

CHAPTER 5   CHARACTERISATION OF GLY1ORF1 HOMOLOGS

5.5 Introduction

The Gly1ORF1 protein sequence is very highly conserved in *Neisseria* spp*.* The closest sequence homolog from outside of the genus has been identified in *Haemophilus sputorum* (51% identity) The PSI-BLAST algorithm allows the identification of distant relatives of the protein by using a custom, position-specific, scoring matrix (Altschul *et al.*, 1997). The database search done previously by Sathyamurthy (Sathyamurthy, 2011) identified many interesting proteins encoded in the genomes of Proteo- and Cyanobacteria which colonise a wide range of organisms, from humans and mammals to fish and even hydra. As the database is constantly updated with new genome sequences, subsequent searches resulted in new, potentially interesting proteins.

This study focuses on proteins encoded by the genomes of *Mannheimia haemolytica,* which is a known cattle pathogen residing in the respiratory tract, and *Haemophilus influenzae, a* common human pathogen that colonises similar niches to *N. meningitidis,* and is also a major causative agent of septicaemia and meningitis.

The main aim of this study was to solve the crystal structures of the selected proteins to check if the novel fold found in the neisserial Gly1ORF1 structure is more widespread in the bacterial world, especially in pathogenic bacteria. This study also investigated if the selected proteins have similar functions to neisserial Gly1ORF1. The proteins were tested for the ability to bind haemin *in vitro* and interact with red blood cells. Finally, initial biophysical characterisation of the HinfGly1 protein, and the neisserial Gly1ORF1 paralog, were performed.

5.6 Results and Discussion

5.6.1 Bioinformatics

In order to identify distant relatives of Gly1ORF1 homologs, the PSI-BLAST algorithm was employed, and set to exclude the neisserial protein (Accession: NP_273818.1) (Altschul *et al.*, 1997). Most of the results were hypothetical proteins from various Gram negative pathogenic bacteria. Several protein sequences were chosen for further analysis. These proteins were chosen based on the presence of an N-terminal signal peptide, and at least one predicted disulphide bond. The proteins are derived from pathogens which possess haemolytic activity and/or are responsible for septicaemia in either humans or animals. The development of animal models of septicaemia caused by pathogens expressing proteins homologous to the neisserial Gly1ORF1 protein would be beneficial in order to understand the role of this protein in neisserial infection. The multiple sequence alignment of several possible Gly1ORF1 homologs identified in pathogenic microorganisms associated with septicaemia is shown in Fig. 5.1. Highly conserved residues (present in 50% or more analysed sequences) were identified on the structural model of Gly1ORF1 K12A structure (Fig. 5.2) The results show that the majority of conserved residues are localised between two β-sheets of a single Gly1ORF1 monomer. This region of the protein might be important for the protein function or play a role in structural integrity of Gly1ORF1 protein family.

The most interesting candidates for further study were proteins found in *Haemophilus influenzae,* a human pathogen that is one of the main causes of septicaemia, and *Mannheimia haemolytica,* a cattle pathogen that is also reported to cause septicaemia in sheep (Mackie *et al.*, 1995).

The search for Gly1ORF1 homologs and paralogs for further study was performed in April 2012. Since then, many new sequences appeared on the database; however, due to the time restrictions they were not investigated. The sequences were selected after five iterations.

```
Neis_meni    1 ----------ADTVESCKTDNNKYIEVQKIN-----RNLYEVSFGSAA-KKEIAIR----
Manh_haem    1 ---------NNNVVYSCTTTDNQTLKVTKE------GGNYVYSHCNTTFKNPVKEA----
Hinf_infl    1 ----------ANNILSCTAENGSPVSVTKN------GSDYEFTYCQVSFKNPVKQV----
Acin_niph    1 ----ACPSQSK-TVFNCTTTNNKVIQVCDA------GNTISVSFGKASAAPELAIT----
Vibr_n418    1 ----ASPSVFANTLFYCQTQSGKQVEVQDF------GSIIHYREGKKLVEAELQLT----
Prot_mira    1 -------NTHTVTVFQGANKNQKQVNISLV------NEQYIVQFGKLNQKPDIIIT----
Ente_cloa    1 ---------NTELVFQGMTKSKKIIKIERT------DEILKCTYGNDKKNDLTLETRLFS
Morg_morg    1 -----------QEITGQLENNKYVSVVVEK-----GKTPASRYCTLA-KPEITLP----
Prov_stua    1 -------NTHTVTVFQGVNKNQKVNISLA------NGQYIVQFGKLNQKPDITMV----
Avib_para    1 ---------EDYIVESCFTAKGKQIALIES------EGYYELSFGKPG-KPELVFK----
Mora_capr    1 -----------ETVFLGDTTNGKQVKLIDN-----GNSFTMEFSKGK-KTELNFS----
Pseu_M47T    1 STPPSNLCLDSVALFQGDTASGKQIALCGSYDAAGELSGLQYRYCRGP-QPELVYP----
Acti_pleu    1 ---------ENISIFSCTAEDGNQILIQKT------GADYQFSYCKILFKNPAKQV----
                 *&@    *   *   *   &  &            &   @  @*@&*     *  *&

Neis_meni   41 NSKADLLGRSDRWQG-MGSGRWATMKFQN----GEFMVTIWTGF-DSVT--HTESSGVVV
Manh_haem   42 LKNP----ASEIAGG--SQFTTISLELRN----AGKSYIVGHIE-ADPKSPFEASVQIQD
Hinf_infl   41 FANQ----DSYVATG--SGFITSSLEMRN----NGTSYTIQFVQ-PHNSNSIEE--PMLY
Acin_niph   46 VPRD--KVTTYQWEG-FGRYENYAINIPN----GKTIYRVNDSL-DKMS--QEYTAGVEV
Vibr_n418   47 VSRD--LASTYQWIG-AGSNEYYSVSIPNPSAQNAIVYRVFTSR-ERKPD-GKMESGIDV
Prot_mira   44 RKPEQLEGRYLNPKGDTGTAQIYEMDFRN----GNYTYTISSSY-FDSK----HEALVDV
Ente_cloa   46 EKIPTIFSSDVQLGNKSGPYVFNTVTFHN----GEYTYSVSALS-DINVPSDETFTGVYV
Morg_morg   39 -VKGNAKNGVYVGQTMFSGGGSAYIRFEN----GPFNYVVYSGM-GRGW----EFTGLAV
Prov_stua   44 RKPEQLAARYLTPKGDRFNGKIYEVDFRN----GDYTYTVSSSY-LDSK----HIAHVDV
Avib_para   41 NKIADVEARSEKWNG-VGS-YWEAYNLKN----GQYYWVVYKSE-NRVSDEHEFSGGVTV
Mora_capr   39 IAKHKIDYTEGAWKAESGYYWTMPYKGNTYTVFSSMINLDENSK-EFSEP--KKEAGVIV
Pseu_M47T   56 AQLKDSLAKFKSNHYFRYQTDYSQISFTS----GSYAGSMYSNYDGEGSPGNPRTAGVTV
Acti_pleu   42 LKNE----GSYVATG--SGFTTSALEMKS----KSHSYTIEFVQ-PRGTNVIDS--PMLY
                *        &   &          &  **      &  @  @*&                 && @

Neis_meni   93 ERRG-KEVARVGGTPKTAQANFNDDDFS---------------
Manh_haem   91 IKTG-NDITSFEGRSDKPIRHNFDRKLMRKSGFAA--------
Hinf_infl   88 ITNG-SKMDTVSGKAGSAT-QNFERRSMKAS------------
Acin_niph   96 SNND-KLLATVEGAANKKITSKIQGIKLRPEM-----------
Vibr_n418  102 LSQG-KIMATILGQ-NESLYEQLIGVDLPAE------------
Prot_mira   95 YQND-KLLTTVNGLPNTIVDNLSEHIFAI--------------
Ente_cloa  101 SGNN-IKTNKIKGIEGTIQDNFESLFTFNSSRNK---------
Morg_morg   89 YKGD-KVINKKVGKTDLGIYDVLNYGLPEDESGEIYAMDPNSQ
Prov_stua   95 YQKD-KLLTSISGLPNTIVDNLHEHIFDLPSDD----------
Avib_para   94 SKGD-KEIATVSG----------ERF-----------------
Mora_capr   96 NQDK-PNSKTILGNPKQIKVDELENYGTPN-------------
Pseu_M47T  112 VSQAGAPVADIAGKAIGKDTLQALTDKVACDADQALGCSQ---
Acti_pleu   89 ITKG-NNTTEIKGDTNKKVGQHFEYSKMNIL------------
                 *  &  *&        &     *
```

Figure 5.1 Multiple sequence alignment of the Gly1ORF1 homologs.

The homologs were identified by PSI-BLAST search using processed neisserial sequence as a query The alignment was done by ClustalW2 (Larkin *et al.*, 2007) and the Boxshade alignment. Residues with 100 % sequence conservation are highlighted in red, with 71-99% - in blue and with 50-70% - in green. The conserved residues based on their properties are marked as follows:'*' - hydrophilic, '&' – hydrophobic and '@' – aromatic residues. Sequences used for the alignment: Neis_meni: *N. meningitidis* NP_273818.1, Manh_haem: *Mannheimia haemolytica* ZP_04977153.1, Hinf_infl: *H. influenzae* R2866 WP_005667853.1, Acin_niph: *Acinetobacter* sp. NIPH 284, ENW83788.1; Vibr_n418: *Vibrio* sp. N418, WP_009388134.1; Prot_mira: *Proteus mirabilis* WGLW6, EKA96180.1; Ente_cloa: *Enterobacter cloacae* P101, AHE73089.1; Morg_morg: *Morganella morganii*, WP_024473070.1; Prov_stua: Providencia stuartii MRSN 2154, AFH94386.1; Avib_para: *Avibacterium paragallinarum*, WP_017807504.1; Mora_capr: *Moraxella caprae*, WP_029102082.1; Pseu_M47T: *Pseudomonas* sp. M47T1, EIK98382.1; Acti_pleu: *Actinobacillus pleuropneumoniae*, WP_005607867.1.
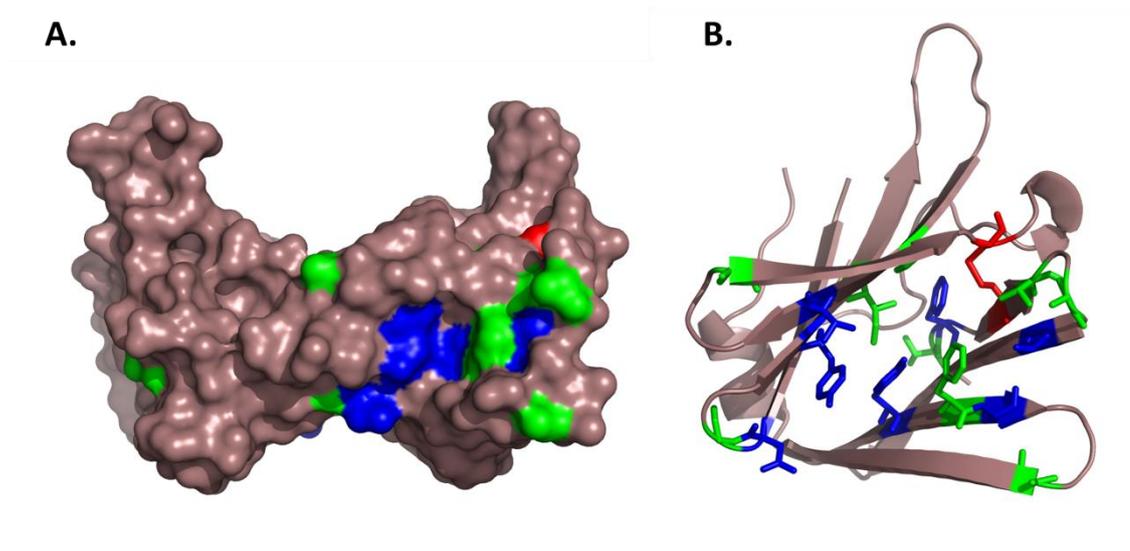
Figure 5.2 Localisation of the conserved residues in the Gly1ORF1 homologs

A. Surface representation of Gly1ORF1 K12A structure with the conserved residues (identity) shown in red; in 71-99% - in blue and 50-70% in green. **B.** A cartoon representation of the Gly1ORF1 K12A monomer structure with conserved residues shown as sticks. The images were generated and rendered using Pymol.

Search of the database with PSI-blast algorithm came out with four interesting hits in *Mannheimia haemolytica*. Three proteins from different strains were almost identical, with a difference in one amino acid only, and one protein being similar to the other proteins only in 41% (MHA0579). The accession numbers and sequence similarity between hits in *Mannheimia haemolytica* are summarised in Table 5.1.

A further search for homologs using the *mh0603* sequence resulted in the identification of a potentially promising protein sequence in the *Mannheimia haemolytica* genome that is predicted to be an adhesin. The protein, referred to as ManhGly1ORF1 paralog or MHA2262, is surprisingly located in close proximity to genes involved in haemin transport, and an outer membrane Fe^{3+} receptor (Fig. 5.3). Using the ManhGly1ORF1 paralog another search with psi-blast showed similarity with the neisserial Gly1ORF1 paralog, a possible adhesin encoded by the *nmb2095* gene.

Table 5.1 Gly1ORF1 homologs and paralogs

	Accession number	Name of the protein and source	Sequence	Sequence similarity with NMB0776	Protein physico-chemical properties and subcellular localisation prediction
1	NP_273818.1	NMB0776 [Neisseria meningitidis MC58]	MKKMFLSAVLLLSAAAQTVWADTVFSCKTDNNKYIEVQKINRNLYEYSFGSAAKKEIAIRNSKADLLGRS DRWQGMGSGRWATMKFQNGEFMYTIWTGFDSVTHTESSGVVVERRGKEVARVGCTPKTAQANFNDDDFS	N/A	Unprocessed protein: 139, MW: 15.6 kDa, pI: 9.14 Processed protein: 119aa, MW: 13.4 kDa, pI: 8.61
2	ZP_04977153.1	hypothetical protein MHA_0579 [Mannheimia haemolytica PHL213]	MKKMAIVALSAFFSMNAFANNNVVYSCTTTDNQTLKVTKEGGNYVYSHGNTTFKNPVKEALKNPASEIAG GSQFTTISLELRNAGKSYIVGHIEADPKSPFEASVQIQDIKTGNDITSFECRSDKPIRHNFDRKLMRKSG FAA	26 identical residues in the processed protein	Unprocessed protein: 143 aa, MW: 15.7 kDa, pI: 9.27 Processed protein: 124 aa, MW: 13.66 kDa, pI: 8.91
3	ZP_05992257.1	hypothetical protein COI_1584 [Mannheimia haemolytica serotype A2 str. OVINE]	MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKKAVTNDGSIIA RGSGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQSKEQPSIGSVDCDTRKKIHYKFDVHLMN TL	20 identical residues over the processed protein	Unprocessed protein: 142 aa, MW: 15.7 kDa, pI: 8.72 Processed protein: 123 aa, MW: 13.6 kDa, pI: 7.95
4	ZP_04977739.1	hypothetical protein MHA_1205 [Mannheimia haemolytica PHL213]	MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKKAVTNDGSIIA RGSGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQSKEQPSIGSVDCDTRKKSHYKFDVHLMN TL	20 identical residues in the processed protein	Unprocessed protein: 142 aa, MW: 15.7 kDa, pI: 8.72 Processed protein: 123 aa, MW: 13.6 kDa, pI: 7.95
5	ABG89179.1	hypothetical protein MH_0603 [Mannheimia haemolytica]	MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKKAVTNDGSIIA RGSGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQRKEQPSIGSVDCDTRKKSHYKFDVHLMNTL	21 identical residues in the processed protein	Unprocessed protein: 142 aa, MW: 15.7 kDa, pI: 8.94 Processed protein: 123 aa, MW: 13.66 kDa, pI: 8.52

	Accession number	Name of the protein and source	Sequence	Sequence similarity with NMB0776	Protein physico-chemical properties and subcellular localisation prediction
6	ZP_04978 753.1	possible adhesin MHA_2262 [Mannheimia haemolytica PHL213]	MKKLMIFATTAMIVSNLAHAAG EQSDAREAHESTVKTSTVKYS CQNGKKLSVKYGFNKQGIPTY AEAKLS GKKRFMPINLYTTDATGTNFG DENNFSLYGDPMEFTNHRKAS VNIQSPASEILYKGCTPQK	18 identical residues in the processed protein	Unprocessed protein: 131 aa, MW: 14.4 kDa, pI: 9.34

Processed protein: 111 aa, MW: 12.26 kDa, pI: 9.08 |
| 7 | WP_0056 67853.1 | hypothetical protein [Haemophilus influenzae] | MKKLLTIGAVAMFATPAFAANN IFSCTAENGSPVSVTKNGSDYE FTYGQVSFKNPVKQVFANQDS YVATG SGFITSSLEMRNNGTSYTIQFV QPHNSNSIEEPMLYITNGSKMD TVSCKAGSATQNFERRSMKAS | 26 identical residues in the processed protein | Unprocessed protein: 135 aa, MW: 14.6 kDa, pI: 8.58

Processed protein: 116 aa, MW: 12.6 kDa, pI: 6.81 |
| 8 | NP_27508 3.1 | adhesin complex protein NMB2095 [Neisseria meningitidis MC58] | MKLLTTAILSSAIALSSMAAAAG TDNPTVAKKTVSYVCQQGKKV KVTYGFNKQGLTTYASAVINGK RVQM PVNLDKSDNVETFYGKEGGYV LGTGVMDGKSYRKQPIMITAP DNQIVFKDCSPR | 15 identcal residues in the processed | Unprocessed protein: 124 aa, MW: 13.3 kDa, pI: 9.6

Processed protein: 103 aa, MW: 11.2 kDa, pI: 9.52 |

The MC58 Gly1ORF1 and a potential homolog in *Haemophilus influenzae* share 26 identical residues in the processed protein. The identified homologs from *Mannheimia haemolytica* share between 20-26 identical residues with the MC58 Gly1ORF1. Gly1ORF1 paralogs in *Mannheimia* and *Neisseria* share with neisserial Gly1ORF1 18 and 15 identical residues respectively, and 42 identical residues between each other. Homologs from *Mannheimia* show quite high similarity with the *H. influenzae* homolog, with 40-41 identical residues when analysed using the Multalin sequence alignment tool, on the default settings (URL: http://multalin.toulouse.inra.fr) (Corpet, 1988).
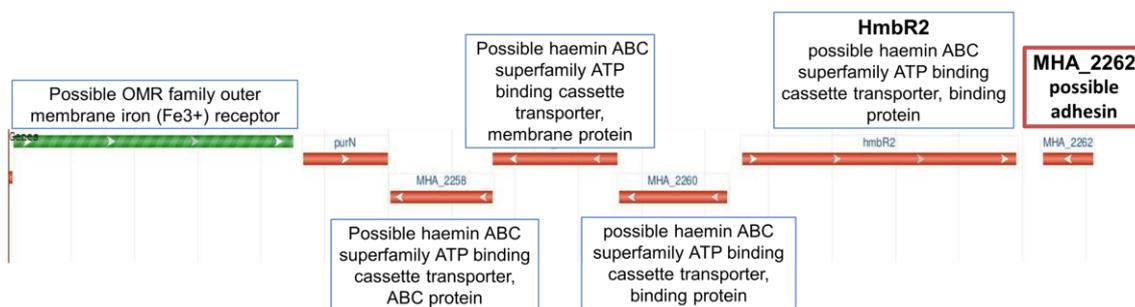
Figure 5.3 *Mha2262* **gene locus in the genome of** *Mannheimia haemolytica***.**

Screen shot of the *Mannheimia haemolytica* genomic sequence (DS264670.1: 10048 to 18800) showing the region that encodes the Gly1ORF1 paralog (*mha2262*). Three haemin ABS transporter genes and outer membrane iron (Fe^{3+}) receptor are located downstream from *mha2262*.

All analysed proteins have a predicted N-terminal signal peptide, and they all contain two cysteine residues that are predicted to form a disulphide bond (Disulfind server: http://disulfind.dsi.unifi.it/; (Ceroni *et al.*, 2006). The secondary structure was predicted using the Sable server (URL: http://sable.cchmc.org/, (Adamczak *et al.*, 2005)), and showed a consensus helical region in the N-terminus that corresponds to the predicted N-terminal signal peptide. The secondary structure predictions have a high content of beta strand regions. The positions of the predicted β-strand regions are very similar, as are the positions of the two cysteine residues. The alignment of the secondary structure predictions for all the proteins shows that the proteins could form two groups. The first group would be comprised of the *Mannheimia* Gly1ORF1 paralog (MHA2262) and neisserial Gly1ORF1 paralog (NMB2095), as the similarities between these two are higher when compared to the other proteins. The second group is comprised of the Gly1ORF1 homologs from *Mannheimia* and *H. influenzae* together with neisserial Gly1ORF1. These proteins have similar positions of the alpha helical domains and strand regions (Fig. 5.2).

All analysed Gly1ORF1 homologs and paralogs were predicted to have an extracellular localisation signal, when analysed by PSORTb version 3.0.2 and

SignalP 4.2. The *H. influenzae* Gly1ORF1 homolog was predicted to be secreted into the extracellular space with high confidence.

The search for motifs using InterProScan did not detect any conserved motifs, apart from a signal peptide for all of the analysed proteins, and a transmembrane region for the *H. influenzae* protein.

For further characterisation of proteins in this study, Hflu (referred to as HinfGly1ORF1), MHA1205 (referred to as ManhGly1ORF1), and NMB2095 (referred to as neisserial Gly1ORF1 paralog) were cloned into expression vectors, expressed and purified.

NMB0776 MKKMFLSAVLLLSAAAQTVWADTVFSCKTDNNKYIEVQKINRNLYEYSFGSAAKKEIAIRNSKADLLGRSDRWQGMGSGRWATMKFQNGEFMYTIWTGFDSVTHTESSGVVVERRGKEVARVGCTPKTAQANFNDDDFS

Hflu MKKLLTIGAVAMFATPAFAANNIFSCTAENGSPVSVTKNGSDYEFTYGQVSFKNPVKQVFANQDSYVATGSGFITSSLEMRNNGTSYTIQFVQPHNSNSIEEPMLYITNGSKMDTVSCKAGSATQNFERRSMKAS

MHA0579 MKKMAIVALSAFFSMNAFANNNVVYSCTTTDNQTLKVTKEGGNYVYSHGNTTFKNPVKEALKNPASEIAGGSQFTTISLELRNAGKSYIVGHIEADPKSPFEASVQIQDIKTGNDITSFECRSDKPIRHNFDRKLMRKSGFAA

MHA1205 MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKKAVTNDGSIIARGSGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQRKEQPSIGSVDCDTRKKSHYKFDVHLMNTL

MHA2262 MKKLMIFATTAMIVSNLAHAAGEQSDAREAHESTVKTSTVKYSCQNGKKLSVKYGFNKQGIPTYAEAKLSGKKRFMPINLYTTDATGTNFGDENNFSLYGDPMEFTNHRKASVNIQSPASEILYKGCTPQK

NMB2095 MKLLTTAILSSAIALSSMAAAAGTDNPTVAKKTVSYVCQQGKKVKVTYGFNKQGLTTYASAVINGKRVQMPVNLDKSDNVETFYGKEGGYVLGTGVMDGKSYRKQPIMITAPDNQIVFKDCSPR

Figure 5.4 Secondary structure prediction for Gly1ORF1 homologs and paralogs using the Sable online tool (Adamczak *et al.*, 2005).

Green arrows indicate predicted β-strand regions, alpha helical regions are in red, and random coil-coil region are in blue.

5.7 Cloning of the Gly1ORF1 homologs and a paralog

The templates for amplification of the Gly1ORF1 homolog sequences consisted of the boiled cells of *H. influenzae* AJ627386 and AM884334 and genomic DNA of *Mannheimia haemolytica.* The strains of *Haemophilus influenzae* were kindly donated by Dr David Wyllie from The University of Oxford. *Mannheimia haemolytica* (50 ng) genomic DNA was obtained from the Intervet Innovation GmbH, Schwabenheim, Germany. The oligonucleotides used for the cloning and PCR conditions are summarised in Table 7.2 in Materials and Methods Section 7.3.2. The PCR resulted in products with the expected sizes (480 and 530 bps respectively) (Fig. 5.5). The neisserial Gly1ORF1 paralog gene (*nmb2095*) was amplified, using the appropriate oligonucleotides, from the genomic DNA of *N. meningitidis* MC58 (Table 7.2).
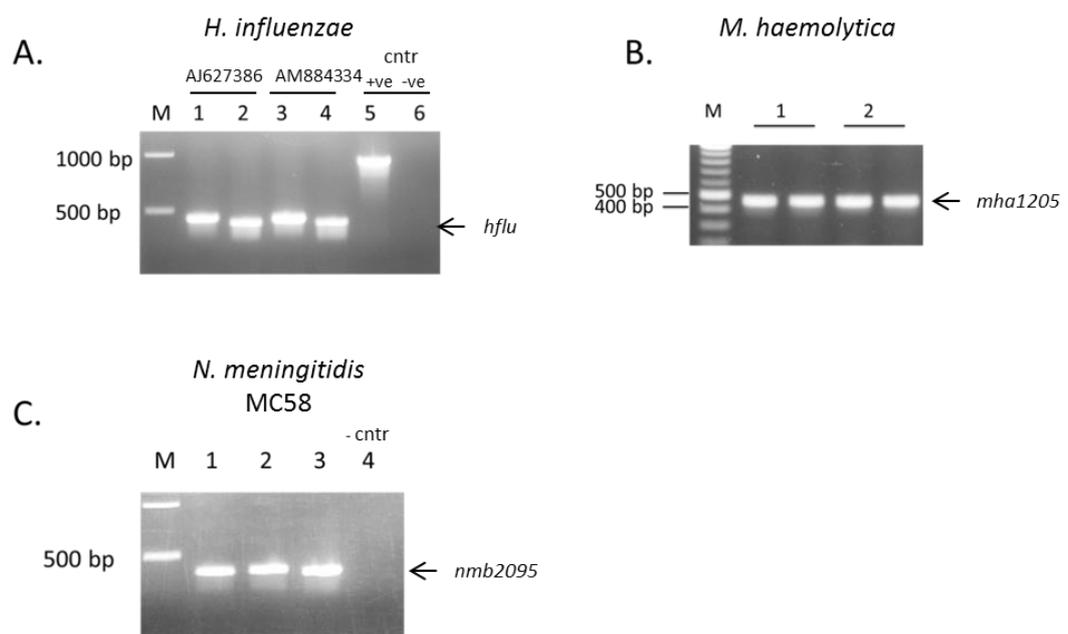
Figure 5.5 PCR products of Gly1ORF1 homologs from *H. influenzae* (Panel. A), *Mannheimia haemolytica* (Panel B.) and neisserial Gly1ORF1 paralog (Panel C.).

A. **Lane 1** - Strain AJ 627386 with oligonucleotides for untagged Gly1ORF1 product, **Lane 2** - AJ 627386 for C-histidine tag product, **Lane 3** – strain AM884334 with primers for untagged product, **Lane 4** – strain AM884334 with primer for C-histidine tag, **Lane 5** - positive control with IgA1 protease β-core primers, **Lane 6** – negative control with Hinf_Gly1 primers with no DNA. **B**. **Lane 1** – ManhGly1ORF1 PCR product from pJONEX4 cloning; **Lane 2** – ManhGly1ORF1 PCR product from cloning into pJONEX4 C-histidine tag; **C.** PCR products of Gly1ORF1 paralog for cloning into: **Lane 1** – pJONEX4, **Lane 2** – pJONEX4 C-histidine tag, **Lane 3** – pET21, **Lane 4** – negative control.

5.7.1 Expression and purification of the HinfGly1 homolog

The HinfGly1ORF1 PCR product, derived from strain AJ627386, was successfully cloned into pJONEX4 and pJONEX4 C-histidine tag vectors and transformed into M72 CaCl$_2$-competent cells as described in Materials and Methods Section 7.3.5. Analysis of the small scale heat-induction of the two constructs confirmed expression of the recombinant proteins in the cell pellet (Fig. 5.6). Analysis of the induced culture supernatants concentrated five times by acetone precipitation did not show presence of recombinant HinfGly1ORF1 (data not shown).
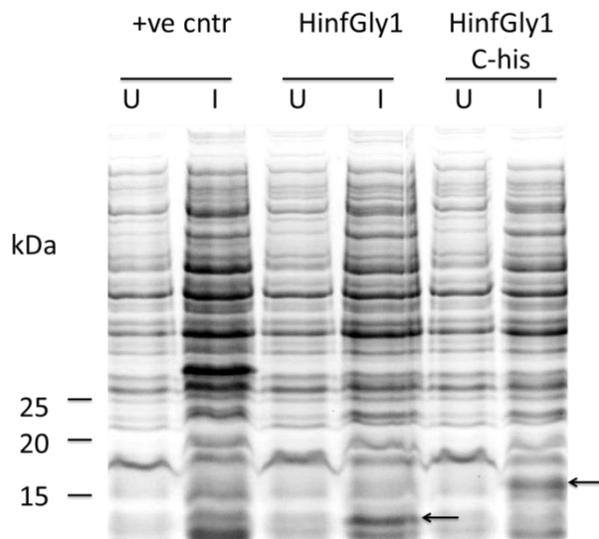
Figure 5.6 Small scale heat induction of HinfGly1ORF1 with and without histidine-tag

Lane U – un-induced samples; **Lane I** – samples after induction. Cells carrying pJONEX4:T5FENΔ19 were used as a positive control for induction (the strong band running above at ~30 kDa). Arrows indicate the induced bands corresponding to the untagged (~13 kDa) and tagged (~17 kDa) version of the protein.

The HinfGly1ORF1 C-histidine tagged protein was expressed on a large scale in a 3 litre fermenter by heat induction as described in Materials and Methods Section 7.5.2. The protein was purified from the soluble fraction of the cell lysate on a Ni-affinity column, in the linear gradient of the imidazole concentration (see Materials and Methods Section 7.6.2.2. Fractions containing HinfGly1ORF1 were pooled together, dialysed against KP buffer, pH 6.5 and further purified by ion exchange chromatography on SP column (see Section 7.6.4.2) This purification step enabled the separation of unprocessed and processed protein, and purification from other contaminants. Processed

HinfGly1ORF1 C-histidine tag was N-terminally sequenced by Professor Arthur Moir, and confirmed the signal peptide predictions. The processed form of HinfGly1ORF1 C-histidine tag starts with the sequence: 'ANNIFSCT'. A summary of the 3 litre induction and purification of HinfGly1ORF1 C-histidine tagged protein is shown in Fig. 5.7.
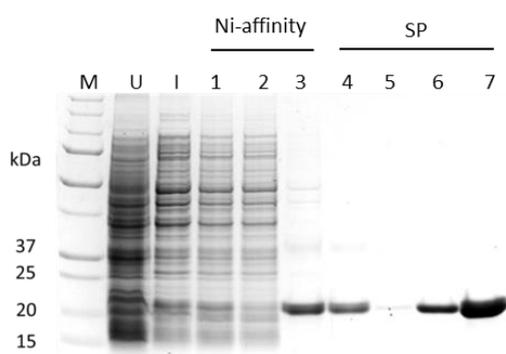
Figure 5.7 Summary of expression and purification of HinfGly1ORF1 C-histidine tagged protein.

Lane M – protein marker; **Lane U** – un-induced cell pellet of M72-pJONEX4C-HinfGly1 with histidine-tag; **Lane I** – induced cell pellet; **Lane 3** – soluble fraction after cell lysis loaded onto Ni-affinity column; **Lane 4** – flow-through from the Ni-affinity column; **Lane 5** – pooled fractions from the Ni-affinity column; **Lane 6** – pooled fractions from the Ni-affinity column after dialysis and centrifugation. This was loaded onto the SP column; **Lane 7** – flow-through from the SP column; **Lane 8** – pooled fractions 1-9 from the SP column; **Lane 9** – fraction 10 from the SP column.

5.7.2 ManhGly1ORF1 (MHA1205) expression and purification

The ManhGly1ORF1 C-histidine tagged DNA fragment was successfully cloned into the pJONEX4 C-histidine tagged vector, and transformed into M72 CaCl$_2$-competent cells (see Materials and Methods Section 7.3). The protein was produced on a large scale by heat induction (see Materials and Methods Section 7.5.2). Analysis of the cell pellets and culture supernatants confirmed successful induction of the protein production. The cell pellets contained mostly insoluble protein Fig. 5.8).

As the objective was the purification of the processed protein, the soluble fraction was passed through the Ni--affinity column, resulting in the purification of the processed protein as described in Materials and Methods Section 7.6.2.2 (Fig. 5.9).
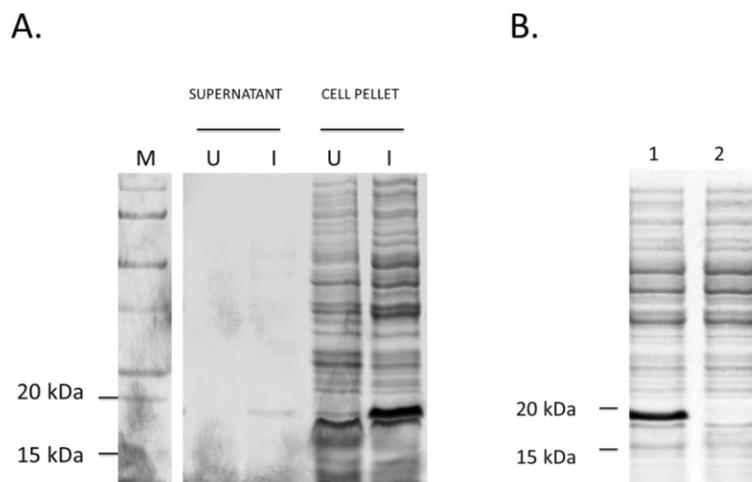
Figure 5.8 Analysis of the large scale expression and solubility of the ManhGly1ORF1 C-histidine tagged protein (MHA1205).

A. Analysis of the cell pellets and culture supernatants before induction (**Lanes U**) and after induction (**Lanes I**) in the 5 litre fermenter culture. **B.** SDS-PAGE analysis of the solubility of the expressed ManhGly1ORF1 C-histidine tagged protein in the cell lysate. **Lane 1** – the whole cell lysate; **Lane 2** – the soluble fraction of the cell lysate.

Attempts to purify ManhGly1ORF1 from the crude supernatant using a number of different types of resins (Ni--affinity, Q-Sepharose, SP-Sepharose, and Phenyl Sepharose) were unsuccessful. However, after precipitation of the proteins present in the supernatant with 3 M ammonium sulphate and dialysis of the precipitated proteins, the protein was successfully purified on the Ni-affinity column as described in Materials and Methods Section 7.6.1.

Comparison of the protein purified from soluble fraction of cell pellet and culture supernatant was performed using SDS-PAGE, and showed that both proteins run at the same size (Fig. 5.9.B). This confirms that ManhGlyORF1 homolog with C-histidine tag in the soluble fraction must be processed by a signal peptidase as it is the same size as the secreted protein purified from the supernatant. Purified ManhGly1ORF1 was used for both crystallisation trials and for the production of rabbit polyclonal antibodies.
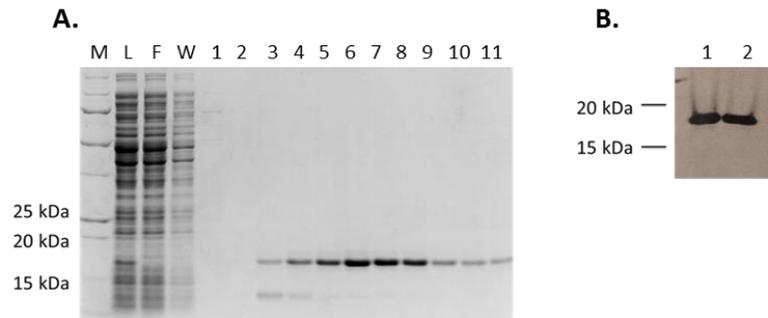
Figure 5.9 Purification of ManhGly1ORF1 homolog on Ni-affinity column

A. Purification from the soluble fraction of the cell lysate. **Lane L** – The soluble fraction of the cell lysate loaded onto the Ni-affinity column; **Lane F** – flow-through from the column; **Lane W** – wash of the column; **Lanes 1 – 11** – elution fractions in a linear gradient of imidazole. **B.** Comparison of ManhGly1ORF1 purified from the soluble fraction of the cell lysate (**Lane 1**), and from the ammonium sulphate precipitated supernatant (**Lane 2**).

5.7.3 Expression and purification of neisserial Gly1ORF1 paralog

The neisserial Gly1ORF1 paralog was successfully cloned into pJONEX4, with and without the C-histidine tag, and into the pET21 expression vector. Small scale expression of the construct showed best results for the auto-induction culture in BL21DE3:pET21 expression system (Fig. 5.10, Panel B.). The neisserial Gly1ORF1 paralog was mostly present in the soluble fraction (data not shown).
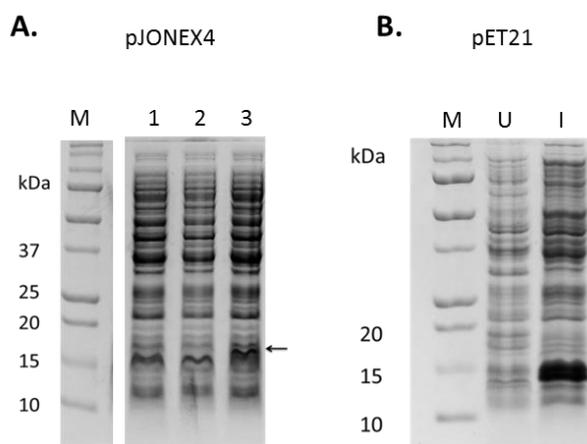
Figure 5.10 SDS-PAGE analysis of the expression of the neisserial Gly1ORF1 paralog in pJONEX4 (Panel A) and pET21-(Panel B) expression vectors

A. Small scale heat induction of M72 cells with following vectors: **Lane 1** – induced empty pJONEX4; **Lane 2** – neisserial Gly1ORF1 paralog in pJONEX4; **Lane 3** – neisserial Gly1ORF1 paralog in pJONEX4 C-histidine tag. **B.** Expression of neisserial Gly1ORF1 paralog in pET21 expression system by auto-induction in BL321DE3 cells. **Lane U** – un-induced sample; **Lane I** – sample after induction. An arrow indicates a protein corresponding to neisserial Gly1ORF1 paralog. Lane M contains a protein marker.

Attempts to purify the protein from the cell lysate using several steps of ion exchange chromatography were unsuccessful (data not shown). A different approach, involving purifying the protein from the periplasmic fraction was undertaken according to the protocol described in Materials and Methods Section 7.6.4.
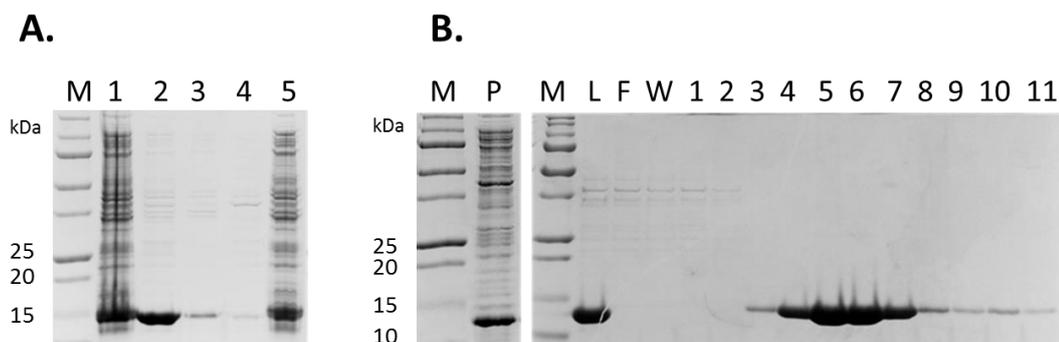
Figure 5.11 Summary of the purification of the neisserial Gly1ORF1 paralog (NMB2095).

A. A protocol for an isolation of periplasmic proteins was used to purify neisserial Gly1ORF1 paralog. Cells after the protein induction were washed with an EDTA-wash buffer, followed by the incubation with the sucrose solution and an osmotic shock. Cells were centrifuged after each step and fractions were analysed by SDS-PAGE as follows: **Lane M** – protein marker; **Lane 1** – whole cell suspension; **Lane 2** –EDTA wash; **Lane 3** – sucrose fraction; **Lane 4** – osmotic shock fraction; **Lane 5** – cell pellet after periplasmic prep. **B**. The EDTA-wash fraction was treated with ammonium sulphate (3.5 M). Prepitated proteins containing contaminants (**Lane P**) were separated from the soluble Gly1ORF1 paralog (**Lane L**) by cenrifugation. Soluble protein fraction was dialysed and further purified using SP column and analysed as follows: **Lane F** – flow-through from the SP column; **Lane W** – wash from the SP column; **Lanes 1-11** – elution fraction in a linear gradient of the NaCl concentration, ranging from 50 mM to 1 M.

Gly1ORF1 paralog was mostly present in the first wash fraction with 5 mM EDTA and was >95% pure (Fig. 5.11, Panel A.). Further steps of the protocol for periplasmic purification of protein including osmotic shock did not produce significantly more protein. The wash fraction was very viscous, therefore DNA contamination was measured calculating the ratio of absorbance at 260 nm to 280 nm; the ratio exceeded 1.4, which indicates high amount of DNA. Thus, polyethylenimine (PEI) precipitation of DNA was performed (see Materials and Methods section 7.6.4.1). The high solubility of the protein enabled further

purification by precipitation of the contaminating bands with 3.5 M ammonium sulphate. Final purification of the protein was achieved by performing the cation exchange chromatography using the HiTrap SP column in KP buffer, pH 7 as described in Materials and Methods Section 7.6.4.2. The summary of the purification is shown in Figure 5.11.

5.7.4 Crystallisation trials for the Gly1ORF1 homologs and the neisserial Gly1ORF1 paralog

All proteins were purified, buffer exchanged into 10 mM Tris-HCl, pH 8, and concentrated up to 30 mg/ml for crystallisation trials at the Department of Molecular Biology and Biotechnology, University of Sheffield. All proteins produced needle-like clusters in many of the screened conditions (Fig. 5.12 – 5.14).
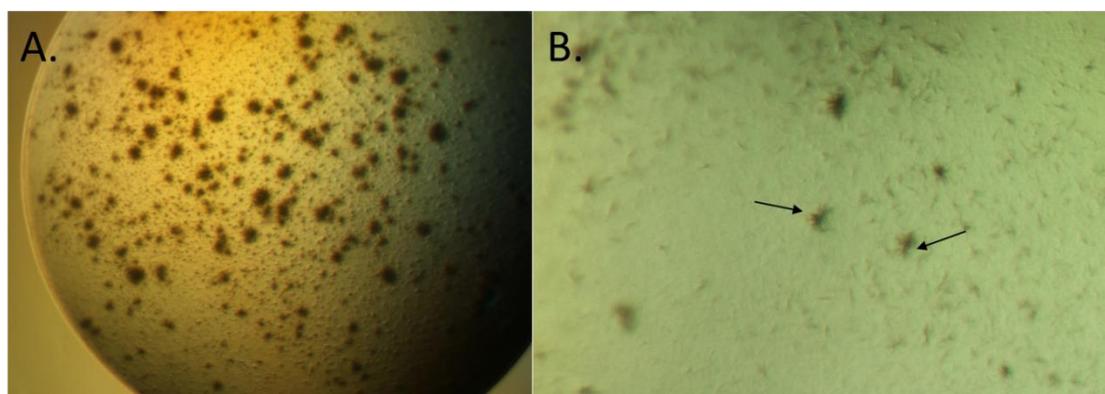
Figure 5.12 Crystallisation trials of the *H. influenzae* GLY1ORF1 homolog.

Proteins crystallised in the Classics Suite, with the sitting drop method, after two weeks incubation at 17°C. A. Needle clusters observed in well H7 in the buffer containing 0.2 M ammonium sulphate and 30% (w/v) PEG 4000. **B**. Needle clusters in F12 conditions with 0.2 M ammonium sulphate and 30% (w/v) PEG 8000.
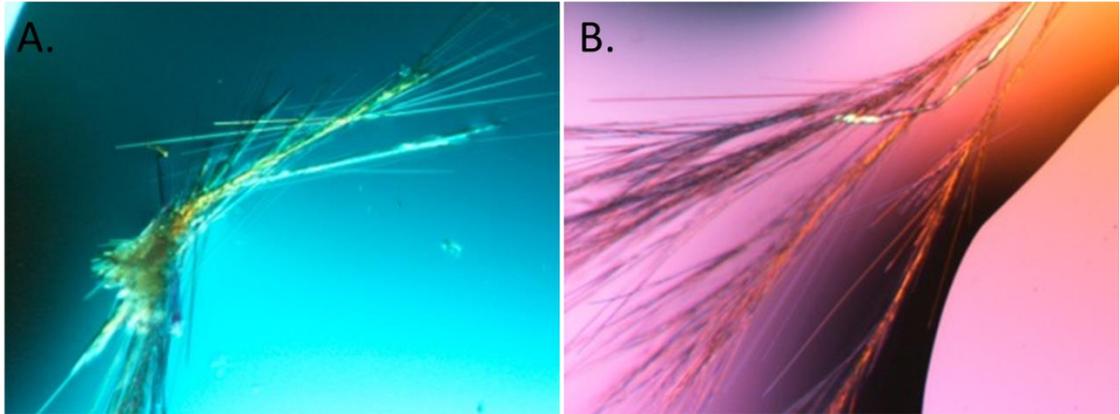
Figure 5.13 Crystals of the ManhGly1ORF1 homolog.

Needle clusters obtained using hanging drop method in: **A.** 0.1 M SPG buffer, pH 8 and 29% (w/v) PEG1500; **B.** in 0.2 M potassium fluoride and 26% (w/v) PEG3350.
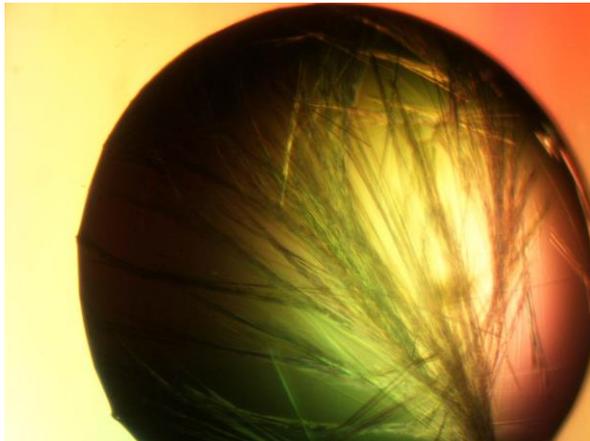
Figure 5.14 Crystals of neisserial Gly1ORF1 paralog (NMB2095).

Crystals were obtained with 40 mg/ml of protein using the sitting drop method in 0.1 M potassium thiocyanate and 30 % (w/v) PEG MME 2000.

HinfGly1ORF1 optimisation trials did not produce crystals suitable for X-ray diffraction, due to their fragile nature. ManhGly1ORF1 did produce crystals at PACT G3 (Qiagen) 96 well screen condition that was big enough for X-ray diffraction. This enabled the collection of a full native data set to 2.6 Å, with a 5.1 multiplicity that was 98.7% complete; however, the data were not good enough to solve the crystal structure of the protein. ManhGly1ORF1 shares little sequence homology with the neisserial Gly1ORF1. This prevented the use of the neisserial Gly1ORF1 structure for molecular replacement. Further optimisation results were unsuccessful.

Attempts to solve the structure of the neisserial Gly1ORF1 paralog (NMB2095) were unsuccessful also due to the very fragile nature of the crystals that were produced.

5.7.5 Characterisation of the Gly1ORF1 homologs and paralog

5.7.5.1 Haemin binding

Purified Gly1ORF1 homologs and paralog were tested for the ability to bind to bovine haemin agarose beads in a qualitative assay (see Materials and Methods Section 7.8.1). Neisserial Gly1ORF1 was selectively removed from the solution and immobilised on the beads. Similarly, ManhGly1ORF1 was also completely removed from the solution. Only weak binding was observed for HinfGly1ORF1 (Fig. 5.15). The neisserial Gly1ORF1 paralog did not appear to bind to the haemin-agarose beads (Fig. 5.16).
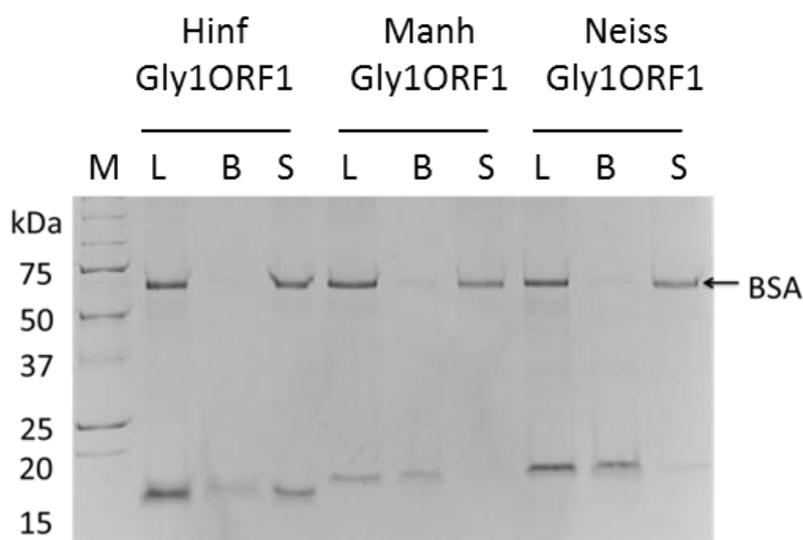
Figure 5.15 Haemin agarose pull down with Gly1ORF1 homologs and the neisserial Gly1ORF1

Lane **L** – samples before incubation with agarose beads; Lane **B** – boiled haemin agarose beads – protein bound to the beads; Lane **S** – supernatant containing unbound proteins. An arrow shows bands corresponding to BSA.
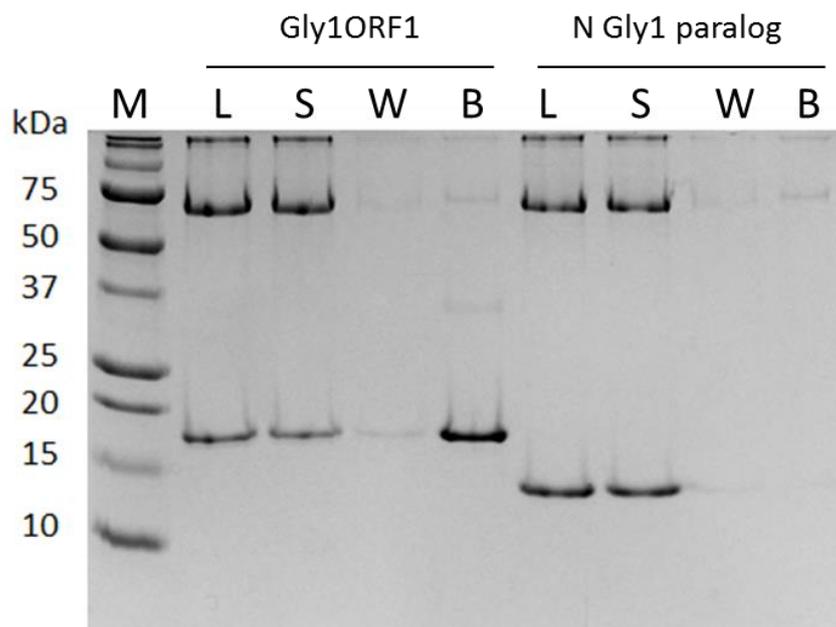
Figure 5.16 Haemin agarose pull down of Gly1ORF1 and the neisserial Gly1ORF1 paralog.

Lane L – samples before incubating with beads, **Lane S** – supernatant containing unbound proteins; **Lane W** – sample containing wash with PBS with 0.5 M NaCl; **Lane B** – boiled haemin agarose beads with bound proteins.

Another approach to test if the purified proteins were able to bind haemin involved haem and Congo red staining of proteins that were separated by denaturing SDS-PAGE, then refolded in the gel as described in Materials and Methods Section 7.10.2. Congo red binding correlates with haem binding activity as described previously (Daskaleros and Payne, 1987; Deneer and Potter, 1989; Smalley *et al.*, 1995). Refolding of the proteins was achieved by subsequent washes of the SDS-PAGE gel with TBG buffer to remove SDS. The wash steps were followed by staining with haemin or Congo red solution dissolved in TBG, and de-staining in TBG buffer. The results are shown in Figure 5.17.

The paralog shows some staining with Congo red and weaker staining with haemin. HinfGly1ORF1 was not detected by any of the staining steps. This experiment again confirms that neisserial Gly1ORF1 can interact with both haemin and Congo red. Although the Gly1ORF1 paralog did not bind to haemin-agarose beads, it showed some evidence of binding to haemin and Congo red

when refolded in-gel. It is possible that steric hindrance, imposed by the immobilisation of haemin on the agarose beads by its carboxyl groups, affects the availability of haemin for this protein, whereas free haemin in solution did not show those limitations. HinfGly1ORF1 did not appear to show binding of either Congo red or haemin. However, this could be a result of incorrect refolding of the protein in the gel and needs to be investigated further.

A.

C.

B.

Figure 5.17 Haem and Congo red binding by proteins refolded in SDS-PAGE gel.

A. Staining with Congo red; **B.** staining with haemin solution; **C.** Coomasie stained SDS-PAGE gel used as a reference. Each lane contained 1.5 µg of protein. Lane 1 – Gly1ORF1; Lane 2 – Gly1ORF1 paralog; Lane 3 – HinfGly1ORF1 paralog.
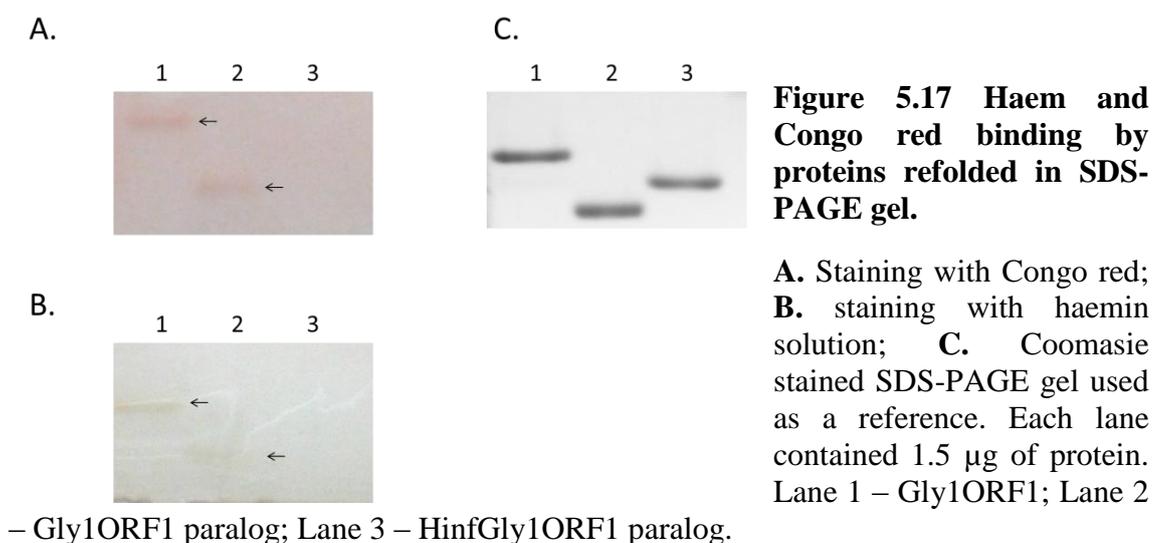
5.7.5.2 Haemin spectra of the Gly1ORF1 homologs and paralog

Gly1ORF1 homologs were analysed for their ability to cause changes in haem absorbance spectra as described in Materials and Methods Section 7.8.2. Results are shown in Figure 5.16. ManhGly1ORF1 caused a slight, but detectable, shift from 385nm to 390 nm (Δλ of 5 nm), which confirms the results obtained in haemin-agarose pull down experiment shown in Fig. 5.15.

HinfGly1ORF1 and the Gly1ORF1 paralog did not cause a detectable shift in the maximum absorbance wavelength. However, changes in the shape of the curves were observed, indicating that there might be some interaction with haemin. It is possible that a higher concentration of the protein could show a more obvious shift in absorbance for ManhGly1ORF1; however, at higher concentrations the protein tends to precipitate out of solution in the presence of

haem. Initial experiments with higher concentrations of the HinfGly1ORF1 protein and the Gly1ORF1 paralog gave similar results (data not shown)
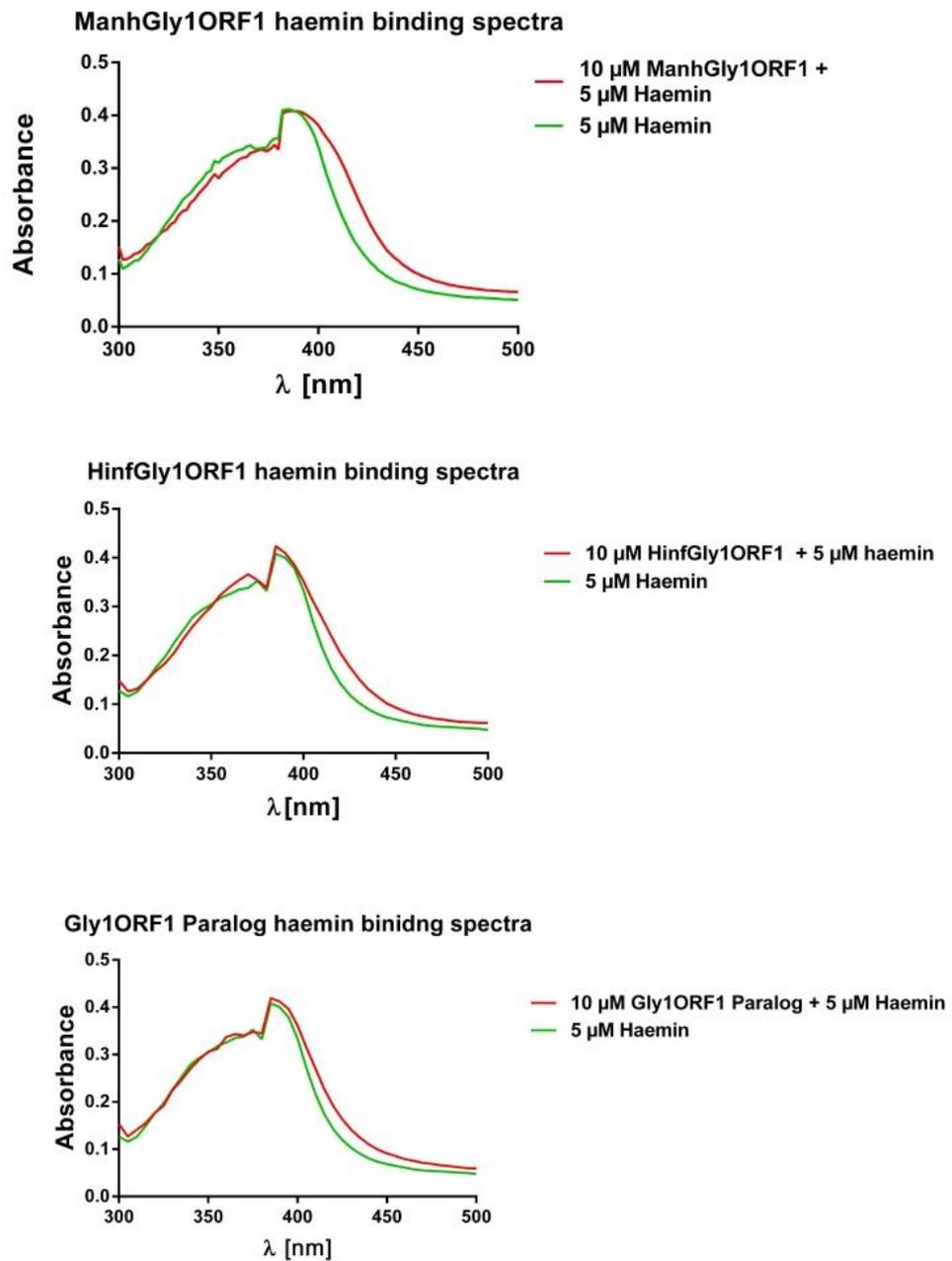
Figure 5.18 Haemin absorbance spectra in the presence and absence of the Gly1ORF1 homologs.

5.7.5.3 Oligomerisation of the Gly1ORF1 homologs and paralog

The crystal structure of Gly1ORF1 shows dimer formation. However, the formation of higher molecular weight complexes was also observed when analysed on SDS PAGE in non-reducing conditions. It was hypothesised that HinfGly1ORF1 and the Gly1ORF1 paralog may also form dimeric or more complex multimeric structures. Analysis of the purified proteins by SDS-PAGE in non-reducing and reducing conditions showed the presence of a higher molecular weight band the Gly1ORF1 paralog at the size of a predicted dimer (Fig. 5.19, Panel A.). The dimeric forms of the Gly1ORF1 paralog were mostly removed from the sample by centrifugation, indicating that dimeric forms, which may be the result of intramolecular disulphide bond formation, might be more prone to aggregation (Fig. 5.20, Panel B.). The formation of multimeric forms was observed after treatment of the protein with glutaraldehyde (see Materials and Methods Section 7.8.5) for both analysed proteins (Fig. 5.19, Panel C).
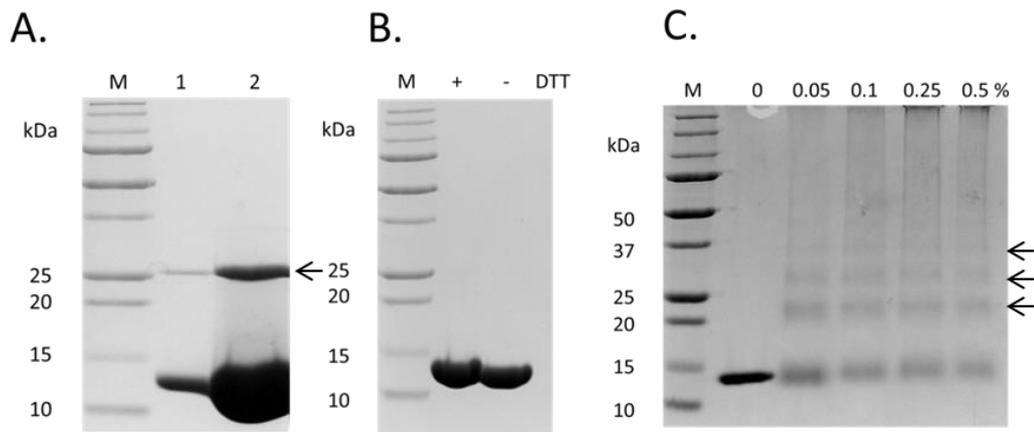
Figure 5.19 Multimerisation of neisserial Gly1ORF1 paralog (NMB2095).

A. Natural formation of the insoluble dimer (indicated with an arrow) – SDS-PAGE in non-reducing conditions of the sample before centrifugation. Lane 1 and 2 show the neisserial paralog at 4 µg per lane and 200 µg per lane respectively. **B.** The soluble fraction after pelleting the precipitant, run in non-reducing (-DTT) and reducing (+DTT) conditions. **C.** Multimer formation after treatment with increasing concentrations of glutaraldehyde. (SDS-PAGE performed under reducing conditions). Gly1ORF1 multimers are indicated with arrows.
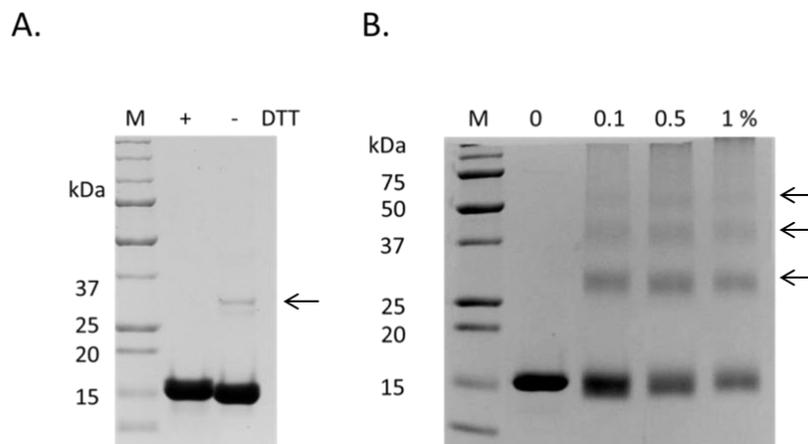
Figure 5.20 Multimerisation of the HinfGly1ORF1 C-histidine tagged protein.

A. Natural dimerisation – SDS-PAGE in reducing (+DTT) and non-reducing (-DTT) conditions (shown with an arrow). **B.** Cross-linking of the HinfGly1ORF1 C-histidine tagged protein with different concentrations of glutaraldehyde. The gel was run in the presence of DTT. Dimers and higher molecular weight oligomers are shown with arrows.

Concentrated HinfGly1ORF1 also shows the formation of dimers and higher molecular weight complexes (Fig. 5.20, Panel B and 5.21, Panel C). Treatment with glutaraldehyde shifts the equilibrium towards higher multimeric forms (5.20, Panel B). Separating the protein using size exclusion chromatography shows three elution peaks, with the main peak corresponding to the dimer when the elution pattern is compared with the known standard Fig. 5.21 Panel A). The second peak likely corresponds to the monomeric form, and a much smaller peak may be a degradation product that cannot be visualised by SDS-PAGE. Analysis of eluted fractions using non-reducing SDS-PAGE shows the presence of faint bands corresponding to molecular weight of the dimer. However, the most intensive band runs at the monomer size (5.21, Panels C. and D.). Analysis of HinfGly1ORF1 migration in the native gel (Fig. 21, Panel B.) before and after size exclusion chromatography does not show any obvious difference in the samples; three populations of bands are observed, with the most intensive band probably being that of the dimer.
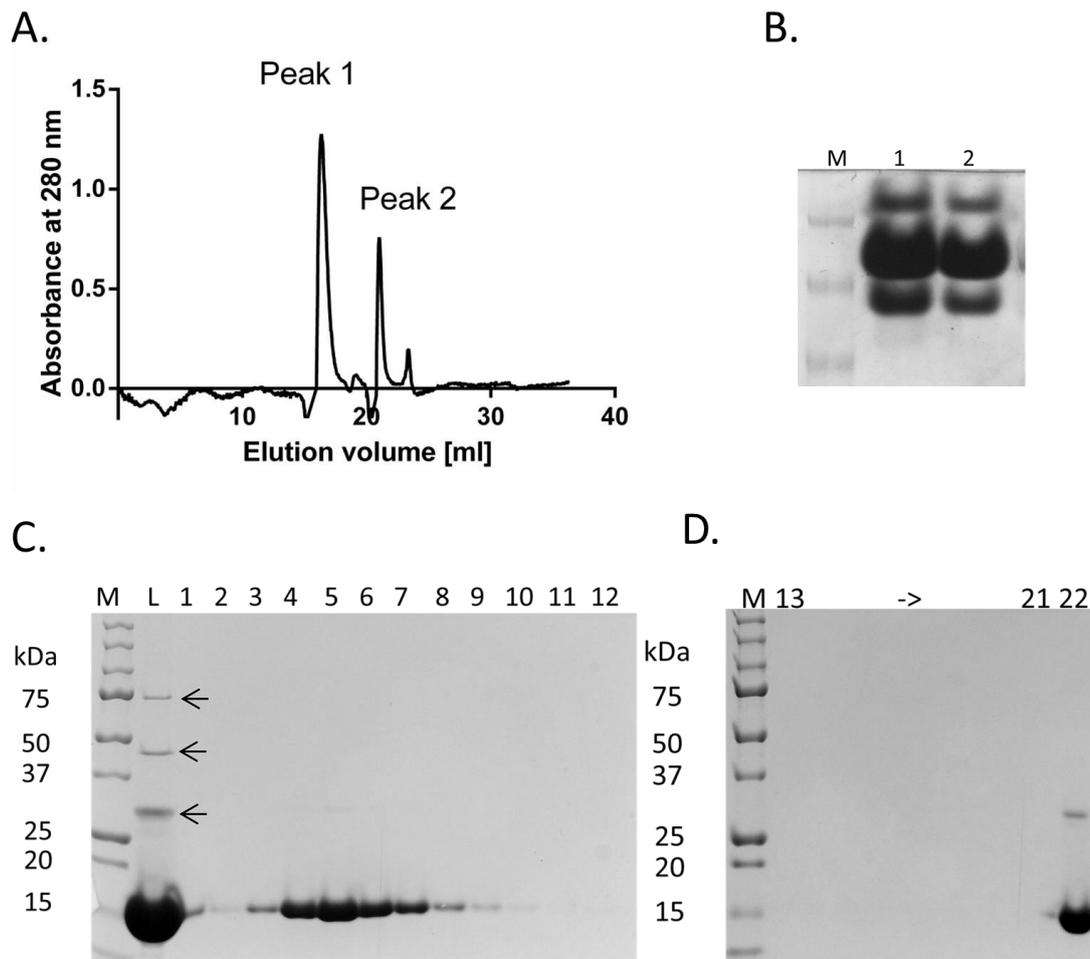
Figure 5.21 Size exclusion chromatography of the HinfGly1 C-histidine tagged protein.

A. UV traces of elution fractions from the SEC column.. **B.** Native gel of the HinfGly1 C-histidine tagged protein. **Lane 1** – the Hinf-Gly1 C-histidine tagged protein before loading on a SEC column. **Lane 2** – Pooled fractions 4-7 from the SEC column. **C.** and **D.** Non-reducing SDS-PAGE analysis of the SEC fractions. **Lane L** – sample loaded on to the column. **Lane 1 – 21** – elution fractions; **Lane 22** – pooled fractions 4-7. Arrows show oligomeric forms of HinfGly1ORF1.

5.7.5.4 CD spectra of HinfGly1ORF1 and the Gly1ORF1 paralog

In order to look at the secondary structure content of the Gly1ORF1 paralog and HinfGly1ORF1, the proteins were analysed by circular dichroism in the Staniforth's laboratory in the Department of Molecular Biology and Biotechnology, University of Sheffield. The spectra for both analysed proteins are shown in Figure 5.22. HinfGly1ORF1 CD spectra of molar elipticity, normalised against the dialysis buffer (PBS), showed a spectrum typical for a

172

protein that consists mainly of beta-sheet structures. Spectra for the neisserial Gly1ORF1 paralog displayed a double dip, which is an indication of a predominantly helical structure within the protein. The data for HinfGly1ORF1 confirm the secondary structure predictions produced by the JPRED3 software shown previously in Fig. 5.4. However, the fact that the neisserial Gly1ORF1 paralog shows a spectrum that might be indicative of a high content of alpha-helical motifs is surprising.
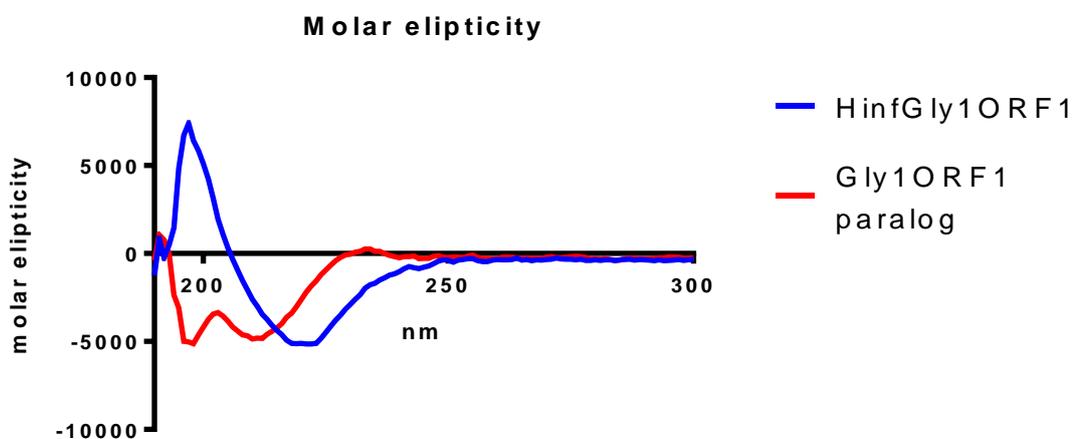
Figure 5.22 CD spectra of HinfGly1ORF1 and the neisserial Gly1ORF paralog, normalised against the dialysis buffer.

5.7.5.5 Binding to red blood cells – FACS analysis

It was previously demonstrated using FACS that Gly1ORF1 is able to bind to red blood cells. This was shown by analysing a shift in the fluorescence of red blood cells upon incubation with FITC-labelled protein (Meadows, 2004). In order to establish if the HinfGly1ORF1 homolog and the neisserial Gly1ORF1 paralog are also able to interact with human red blood cells, both proteins were fluorescently labelled with FITC and incubated with human red blood cells (see Materials and Methods Section 7.9.5). The ability of these proteins to bind to red blood cells were analysed on FACS Callibur.
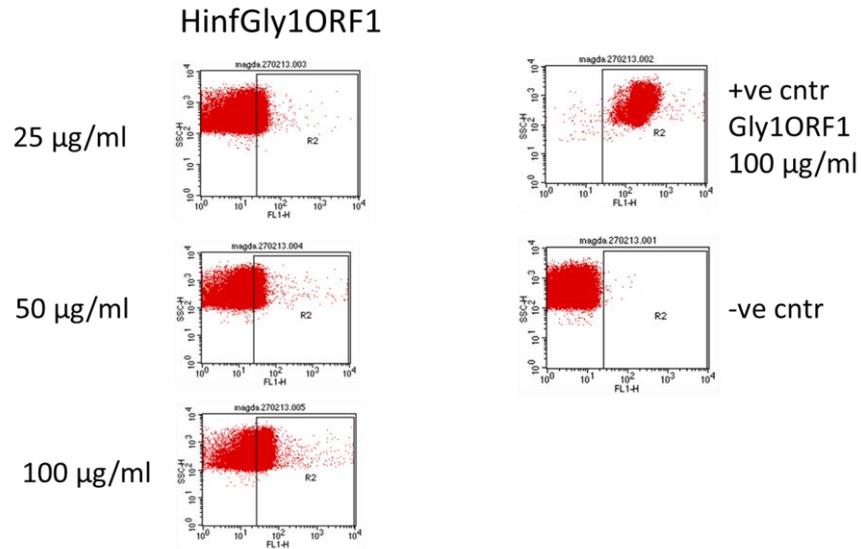
Figure 5.23 FACS analysis of FITC-labelled HinfGly1ORF1 binding to human red blood cells.

A slight dose dependent increase in fluorescence of red blood cells was observed. Neisserial Gly1ORF1 at a concentration of 100 μg/ml was used as a positive control. The negative control shows unstained red blood cells.
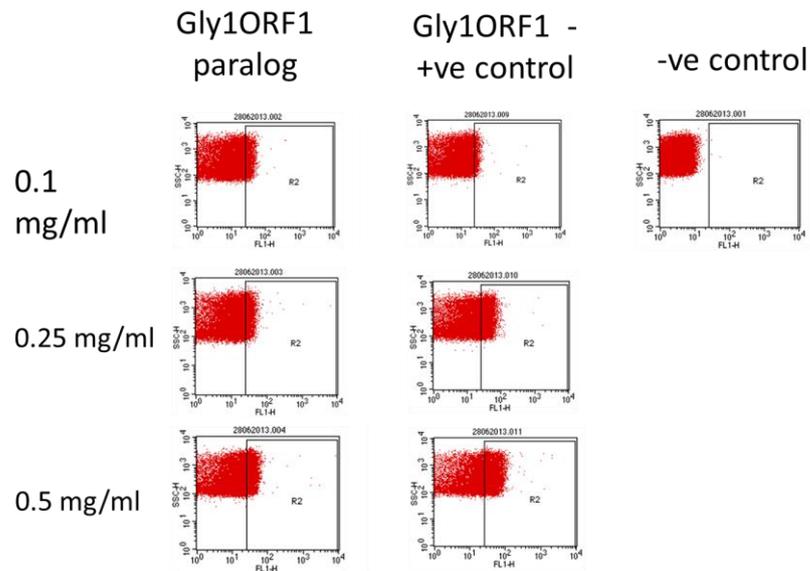
Figure 5.24 FACS analysis of the FITC-labelled neisserial Gly1ORF1 paralog binding to human red blood cells.

The positive control consists of 100 μg/ml of neisserial Gly1ORF1. The negative control consists of unstained red blood cells. No obvious increase in fluorescence was observed with increased dose of the Gly1ORF1 paralog.

HinfGly1ORF1 shows some level of binding to red blood cells, which appears to be dose dependent. The interactions are much weaker than that displayed by Gly1ORF1 (Fig. 5.23). The neisserial Gly1ORF1 paralog shows weak binding to red blood cells (Fig. 5.24). This might be a result of unspecific interactions, as increasing the protein concentration does not increase the shift in fluorescence.

In summary, this chapter aimed to identify and characterise the potential structural homologs of the neisserial Gly1ORF1. The results from a bioinformatics analysis identified a large number of small secreted proteins with conserved positions of cysteine residues and high β-strands content. The largest proportion of hits was from bacteria that are known to be pathogenic to humans or to other animals. Three proteins were identified for further studies: ManhGly1ORF1, HinfGly1ORF1 and the neisserial Gly1ORF1 paralog (NMB2095). These were successfully cloned and expressed in *E.coli*. The main objective of the study was to solve the structure of these proteins in order to see if the bioinformatics predictions correlate with a structural similarity to Gly1ORF1. Although all the proteins did crystallise, they all formed fragile, needle-like clusters, and further optimisation trials proved to be unsuccessful. The fact that all the tested proteins formed similar crystal morphologies might indicate that they share structural features.

The haem-agarose pull-down assay showed detectable binding in the case of the ManhGly1ORF1 protein, with no obvious binding for the other proteins. Although measuring shifts in the maximum absorbance spectra does require further optimisation of the conditions, there was some indication that ManhGly1ORF1 can bind haem. The observation that ManhGly1ORF1 precipitated in the presence of haem, like the neisserial Gly1ORF1 protein does, indicates that these proteins might not only be structurally related, but may also share similar functions. As a result of these findings, a separate, collaborative project was set up with the Department of Food and Rural Affairs (DEFRA) in order to study Gly1ORF1 related proteins in *Mannheimia haemolytica.*

Surprisingly, no obvious binding to haem or Congo red was observed for the HinfGly1ORF1 protein. Although HinfGly1ORF1 shares higher sequence

homology with neisserial Gly1ORF1 than ManhGly1ORF1 does, the main difference is a much lower pI of the mature protein (pI 6.81) compared to the Gly1ORF1-like proteins. One can speculate that this might indicate that the protein, despite having a similar structure, may have a different function.

All tested proteins showed some level of oligomerisation when tested by SDS PAGE in non-reducing conditions. This indicates the presence of intramolecular disulphide bonds. HinfGly1ORF1 was also analysed in a native polyacrylamide gel and separated on a size exclusion column, confirming the formation of dimers and higher molecular weight complexes. Formation of higher molecular weight oligomers was also induced in the presence of glutaraldehyde.

The neisserial Gly1ORF1 paralog seems to be the least similar protein to Gly1ORF1. Differences were detected when comparing the secondary structure predictions of these two proteins. Also, the CD spectra differ from the spectra typical for proteins with high β-strand content. Additionally, there is no evidence of this proteins binding to haem or red blood cells. The neisserial Gly1ORF1 paralog, like the *Mannheimia* GlyORF1 paralog, are predicted to be adhesins, which may explain these differences.

CHAPTER 6   DISCUSSION

Neisseria meningitidis is a successful human commensal and an example of a bacterium with invasive potential. It adapted to human nasopharynx as its stable ecological niche and in most of the cases causes no harm. However, occasionally, the bacteria cross the epithelial barrier and reach the blood stream causing septicaemia and meningitis. The meningococci mastered the art of evasion of the host immune response and exploiting human resources; and can survive and disseminate in the blood stream. The consequences of meningococcal infection can be devastating. The high mortality rate and long-term disability among survivors place *N. meningitidis* in the category of pathogens with high epidemiological and clinical importance. Therefore, there is a great research interest in identifying neisserial virulence factors that would be a basis for the design of novel preventive and therapeutic treatments.

The subject of this study was a small, secreted protein Gly1ORF1 expressed by *Neisseria* spp. The protein is highly conserved in the genus and present in all invasive isolates. The fact that the protein seems to be involved in haem acquisition and an interaction with red blood cells (Meadows, 2004; Sathyamurthy, 2011) suggested that it may be an important, yet uncharacterised neisserial virulence factor.

Gly1ORF1 - structure function study

One of the aims of the study was to determine the crystal structure of the MC58 Gly1ORF1. Previous attempts to crystallise Gly1ORF1 were unsuccessful due to low solubility and aggregation problems. This has been overcome by introducing a single-amino-acid-substitution in the protein. It greatly enhanced the protein solubility enabling successful crystal trials. This mutation also reduced protein precipitation in the presence of haemin. This property of the mutant was of great importance when setting up co-crystallisation trials with haemin. The co-crystallisation trials were also successful.

However, due to technical difficulties, the quality of the collected data for the Gly1ORF1-haemin crystals was not sufficient for solving the structure and identification of the haemin binding site. Further experiments are under way.

The attempts to identify the haem binding site in Gly1ORF1 by other methods were unsuccessful. Mutations previously introduced by Sathyamurthy (Sathyamurthy, 2011) did not inhibit Gly1ORF1 interactions when analysed spectrophotometrically or by the haem-agarose pull-down assay. The mutations were located in the region corresponding to the predicted pocket in the cavity at the dimer's interface. This region was also identified as a potential haem-binding site by several ligand docking servers described in Chapter 2. It cannot be ruled out that single point mutations were not sufficient for the disruption of the haem binding activity of Gly1ORF1. However, it is also plausible that this region despite the prediction is not the actual binding site. Therefore, new mutations were proposed. Many of them were the residues conserved in the Gly1ORF1 homologs. Most of the mutations in this study were targeting residues located in the region between two β-sheets (Fig. 2.2) or in the loops at the predicted dimer interface (Fig. 2.1). None of the mutations disrupted the interactions with the haemin when tested by haemin-agarose pull-down. Also the difficulties in determining the K_d for the wild type Gly1ORF1 were limiting factors in the interpretation of the MST data for three tested mutants. Therefore the attempts should be concentrated on the optimisation of the conditions for determination of the binding activity for the wild-type and the generated mutants.

The K_d data for three Gly1ORF1 mutants gives a clear indication that Gly1ORF1 haem binding affinity is relatively low compared to haemophores in other bacteria. As an example, HasA in *Serratia marcescens* binds haem with the K_d of 18 nM (Izadi *et al.*, 1997). Such low affinity rules out the possibility that Gly1ORF1 could outcompete host haem binding proteins, like HasA does. However, it is possible that it might bind haem in the event of spontaneous lysis and release of haem from haemoglobin, which can be observed in the later stages of meningococcal infection. Haem could be then captured by neisserial haem receptors, such as HmbR by passive transport due their higher affinity to haem than Gly1ORF1 (13 nM and ~300 nM respectively) (Perkins-Balding *et al.*, 2004; Rohde and Dyer, 2004). It is also possible that Gly1ORF1 in presence of some other unknown cofactor or receptor has much higher affinity to the ligand. One such example is HpuA in *N. meningitidis*. This lipoprotein can bind

haemoglobin only when the functional outer membrane receptor HpuB is co-expressed. The interaction between the outer membrane receptor and the lipoprotein is required for efficient binding of the ligand (Rohde and Dyer, 2004). Therefore it would be interesting to see if Gly1ORF1's affinity to haem would also increase in the presence of the HmbR receptor.

The dynamic light scattering experiment confirmed that Gly1ORF1 is present in solution as a mixture of monomeric and oligomeric forms. Although the solved crystal structure shows a dimer formed by two monomers interacting by unknown mechanism (hydrophobic or electrostatic interaction), the experimental data show evidence of formation of inter- and intramolecular disulphide bonds. We hypothesised that the higher molecular weight Gly1ORF1 complexes might be a result of the domain swapping. The reduction of the disulphide bridges linking two β-sheet domains in a monomer of Gly1ORF1 could enable interaction between two reduced monomers and formation of the intra-molecular disulphide links. As a result, the domain-swap dimers and oligomers could be formed. The proposed model for domain-swapping by Gly1ORF1 is demonstrated in Figure 6.1.

Many human plasma proteins have been reported to be controlled by the cleavage of the functional disulphide bonds in the protein backbones. This mechanism has been shown to play an important role in blood homeostasis, thrombosis, blood pressure control and complement regulation (Butera *et al.*, 2014). The reduction of functional disulphide bonds can be mediated by the oxidoreductases circulating in the blood or by thiol/disulphide exchange triggered by conformational change caused by protein-ligand binding or shear of the protein. Therefore, Gly1ORF1 disulphide bonds reduction followed by formation of domain-swapped oligomers could be a mechanism of the modulation of the protein function as a result of protein-host interaction.

The domain-swapping hypothesis could be supported by the fact that mutations of the residues glycine 47 and arginine 48 localised at the dimer interface did not prevent the oligomerisation when analysed by the cross-linking approach. However, the method was not quantitative. Therefore, the mutants could be analysed by dynamic light scattering to study the size distribution of the

Gly1ORF1 oligomers. This would enable detecting subtle differences between mutants in the ability to form oligomers. However, the limited amount of the purified proteins at the time of this study prevented performing the analysis.
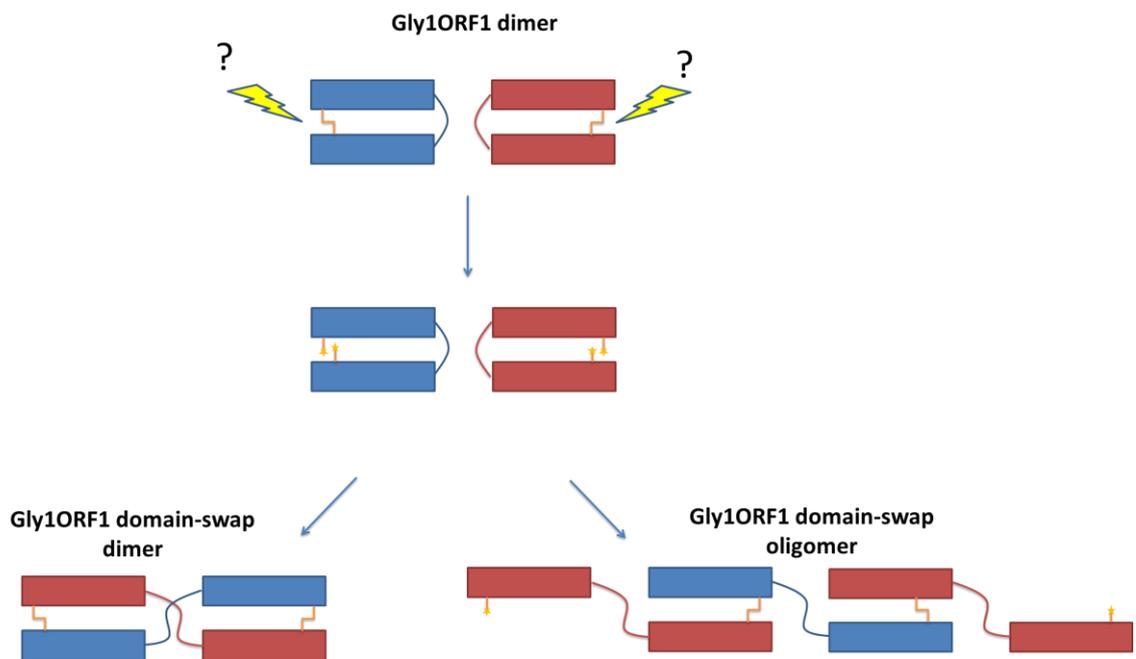
Figure 6.1 Proposed model for domain-swapping in Gly1ORF1.

The rectangles represent two β-sheets domains connected by the linker domain (blue or red line) and a single disulphide bridge (orange line), yellow bolts indicate the oxidoreductases. See description in the text.

Interaction of Gly1ORF1 with red blood cells and EPB4.2

Iron availability is a limiting factor for the growth of *Neisseria* in the blood stream. Iron is almost totally sequestered as a haem complex with haemoglobin and is enclosed within red blood cells. Even during the spontaneous lysis of red blood cells, free haemoglobin and haem are quickly bound by the proteins haptoglobin and haemopoexin, respectively, which circulate in the blood. It is indicative that *Neisseria meningitidis* might have evolved a mechanism that would enable it to acquire the haem from the most abundant source in the blood stream – red blood cells. *Neisseria meningitidis* is not associated with haemolytic activity, apart from few exceptions (verbal communication from Prof.

R. Read). To date, only two potential haemolysins were described in *Neisseria*: FrpA and FrpC. These share sequence homology with the RTX protein family, which is a group of proteins with cytotoxic activity. However, no haemolytic activity mediated by the expression of FrpA and FrpC was observed when these proteins were expressed *in vitro* (Thompson and Sparling, 1993; Thompson *et al.*, 1993).

Therefore, Gly1ORF1 raised research interests when the protein was shown to confer a weak haemolytic activity in *E. coli* (Arvidson *et al.*, 1999). Further experiments performed in our laboratory confirmed that the protein is able to interact with red blood cells causing them to aggregate, and lose their normal discoid shape but the meningococcal protein did not appear to cause haemolysis. In addition, Gly1ORF1 seems to be involved in haem acquisition, as a Gly1ORF1 knock out MC58 strain is unable to grow on haem or haemoglobin as the sole iron source (Sathyamurthy, 2011). The data gathered in our laboratory also suggest that the changes in the morphology of red blood cells can be explained with the interaction of Gly1ORF1 with the EPB4.2 protein – a major component of the RBC's cytoskeleton (Meadows, 2004; Sathyamurthy, 2011).

The protein EPB4.2 accounts for 5% of the erythrocyte membrane content with approximately 2.5×10^5 copies in a single cell (Satchwell *et al.*, 2009). The protein shares sequence homology with the protein family of transglutaminases. However, it lacks the transglutaminase activity due to the absence of a crucial residue in the catalytic site (Korsgren *et al.*, 1990). The EPB4.2 is present in two isoforms, the most prevalent isoform 2 has a molecular weight of 72 kDa and the other, bigger isoform is predicted to be 74 kDa (Sung *et al.*, 1990). The protein EPB4.2 plays a role of the complex linker in the cytoskeleton of the red blood cells. It is involved in maintaining the discoid shape and elasticity of the red blood cells through the interaction with other components of the cytoskeleton. It has been shown that the EPB4.2 is anchored to the plasma membrane by the band 3 protein – a multispanning membrane protein involved in ion transport across the membrane (Bruce, 2008). EPB4.2 also interacts with a transmembrane spanning protein CD47 belonging to the RH complex of proteins (Dahl *et al.*, 2004). From the cytoplasmic site of the membrane,

EPB4.2 binds the protein 4.1, spectrin and ankyrin – proteins responsible for membrane integrity and the characteristic erythrocyte shape. Mutations in the *EPB4.2* gene are associated with the hereditary spherocytosis (Satchwell *et al.*, 2009).

Although no crystal structure of the EPB4.2 is available, a homology model was proposed by Toye and co-workers (Toye *et al.*, 2005) (Fig. 6.2). According to this model, the Gly1ORF1 binding site identified by Meadows (corresponding to residues 174-225) is located in the core Domain-1 of the EPB4.2 protein. The Gly1ORF1 binding site encompasses the asparagine residue at the position 175, which is located in the hinge region between Domain-1 and 2. The variant of the protein with the single point mutation of this residue to tyrosine is described as P4.2 Komatsu and is also manifested as a hereditary spherocytosis (Kanzaki *et al.*, 1995).
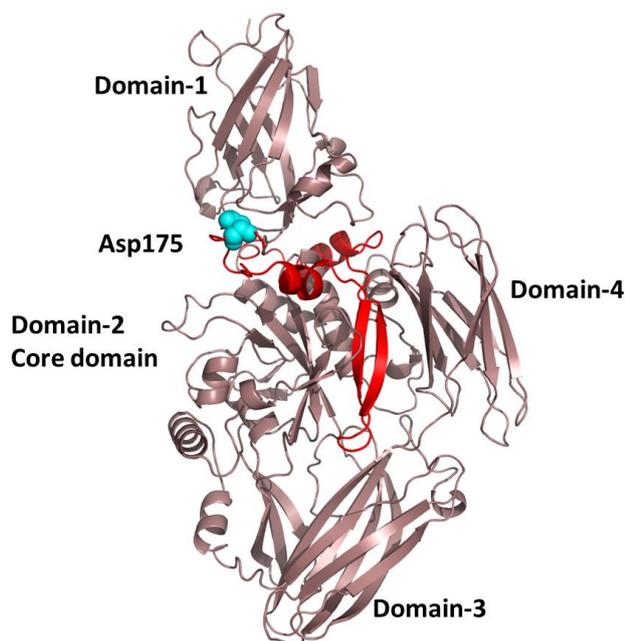
Figure 6.2 Protein EPB4.2 homology model.

Cartoon representation of the EPB4.2 homology model produced using Phyre2 server (Kelley and Sternberg, 2009). The region in red corresponds to residues 174-225 identified as a Gly1ORF1 binding site (Meadows, 2004). The asparagine 175 that is mutated in the Komatsu variant of EPB4.2 is represented as cyan spheres. The image was generated and rendered using Pymol.

The Gly1ORF1 binding region overlaps also with the band 3 protein binding region (residues 187-211) proposed by Bhattacharyya (Bhattacharyya *et al.*, 1999) and the ankyrin binding site (residues 187-200) (Su *et al.*, 2006). The

fact, that Gly1ORF1 has been shown to bind the EPB4.2 region also involved in the interaction with other crucial cytoskeleton proteins might explain the morphological changes in the red blood cells when incubated with Gly1ORF1 observed by Sathyamurthy (Sathyamurthy, 2011)

This study aimed to further investigate the interactions between Gly1ORF1 and the EPB4.2 protein and measure the affinity of the interactions. It also aimed to co-crystallise both proteins to confirm the binding sites identified by Meadows. In order to achieve this, a codon optimised gene was custom-synthesised, sub-cloned and expressed in *E. coli.* The yields of the protein were too low for the crystallisation trials. In addition to this the protein was prone to degradation. However, sufficient protein was purified for the initial test of the interactions with Gly1ORF1 using the BLItz (biolayer interferometry) machine (ForteBio).

The experiments did not show any interactions between the purified Gly1ORF1 and EPB4.2. This could be explained by improper refolding of the denatured EPB4.2 protein during purification. In addition, the protein expressed in *E. coli* lacks post-translational modifications; the EPB4.2 protein undergoes phosphorylation at many sites. It is also myristylated at the N-terminus and palmitoylated at cysteine 203 (in the region identified as a Gly1ORF1 binding site) (Das *et al.*, 1994). It is unlikely that lack of myristoylation would affect the interaction of the protein with Gly1ORF1 as it is far from the predicted Gly1ORF1 binding site. However, the absence of the palmitoyl group on the cysteine 203 might have much more far reaching consequences for the protein-protein interactions. It has been previously reported that the EPB4.2 devoid of the palmitoyl group exhibits greatly reduced incorporation into red blood cell membranes. It has also been suggested that palmitoylation might be important for the interaction with the band 3 protein (Bhattacharyya *et al.*, 1999). Therefore, it is conceivable that the Gly1ORF1 recognising similar region of the protein might display lower affinity to the EPB4.2 protein lacking this group.

Thus, to study protein-protein interactions, the EPB4.2 protein should ideally be purified from the eukaryotic cells to ensure the presence of post-translational modifications. One possible solution is to purify the protein directly from the red blood cell membranes. Our attempts to purify the protein from this source failed

(data not shown). However, the methodology for the purification of EPB4.2 from RBCs employs very harsh conditions, such as high pH or the presence of the chaotropic reagents (Friedrichs *et al.*, 1989); these could also affect protein function.

In future studies, a novel approach avoiding protein purification could be used for studying the EPB4.2 binding affinity. Recently, Khavrutskii and co-workers published a protein purification-free method for measuring kinetics of protein interaction using the microscale thermophoresis approach (Khavrutskii *et al.*, 2013). The protein of the interest was expressed as a GFP-fusion protein in a mammalian cell line and the whole cell lysates were used for the affinity studies. This method reduces the risk of the protein degradation and denaturation during protein purification. In addition, it does not require high levels of expression of the target protein. It by-passes the protein labelling step which can also affect protein stability. .In addition, the proteins expressed in the mammalian cell line should undergo the post-translational modifications that might be vital for their function. This method requires only the ligand to be of high purity.

The purification of functional full-length EPB4.2 for the crystallisation trials seem to be very challenging due to its instability. Therefore, to confirm that the region identified by Meadows (Meadows, 2004) was the actual binding site, point mutations might be introduced into the protein and the affinity to Gly1ORF1 could be analysed as described above. The reduction in the affinity of the interaction would be indicative of the identification of the binding site

A possible solution for the generation of enough protein for crystallisation would be to express and purify only a fragment of EPB4.2. The Gly1ORF1 binding site seems to be localised in the groove between Domain-1 and Domain-2 of EPB4.2 according to the homology model shown in Figure 6.2. Therefore, it is plausible that a fragment containing those two domains would maintain structural integrity of the protein and allow co-crystallisation trials with Gly1ORF1. Expression of the smaller fragment of the EPB4.2 might increase the yields of the expressed proteins and also reduce protein degradation.

The EPB4.2 protein is localised on the inner side of the erythrocyte membrane. Therefore, Gly1ORF1 needs to enter the red blood cells in order to interact with

the EPB4.2 protein. Sathyamurthy suggested that Gly1ORF1 entry inside the red blood cells might be vesicle mediated as Gly1ORF1 has been shown to be incorporated in the OMV when expressed in *E. coli.* Another hypothesis considers interactions of Gly1ORF1 with the erythrocyte membrane lipids leading to a conformational change in the protein. This would enable threading of the protein through the erythrocyte membrane thus accessing the cytoplasm (Sathyamurthy, 2011). Another suggested explanation takes into account involvement of other erythrocyte membrane proteins in the interaction with Gly1ORF1. The proteins could play a role of the receptor for Gly1ORF1 and enable the internalisation of Gly1ORF1 into the cell.

This study tried to establish if Gly1ORF1 can interact with other erythrocyte proteins. The protein was incubated with red blood cells or with the red blood cell ghosts (erythrocytes devoid of haemoglobin) and reversible cross-linkers were added to stabilise transient interactions. The lysed red blood cells and ghosts treated with the detergent were subjected to pull-down experiments using Ni-affinity agarose (data not shown). However, no binding of other than Gly1ORF1 proteins was observed. The western blot analysis also failed to detect any other binding partners at higher than Gly1ORF1 molecular weight – even the band corresponding to the EPB4.2 protein. Interestingly, the analysis of the red blood cell lysates with bound Gly1ORF1 by western blot with anti-Gly1ORF1 antibodies showed the presence of additional proteins slightly smaller than the monomeric and dimeric Gly1ORF1. The proteins were not present in the reducing conditions. It might be indicative of the conformational changes of Gly1ORF1 upon binding to red blood cells affecting protein migration in the non-reducing conditions. This might point towards the hypothesis proposing conformational changes of Gly1ORF1 in the presence of the lipids. It is not clear why similar changes were not observed when Gly1ORF1 was incubated with the ghost membranes.

The future work on this subject could include methods to study structural changes of Gly1ORF1 in the presence of the membrane lipids such as X-ray diffraction and NMR-analysis. The incorporation of the Gly1ORF1 protein in the phospholipid bilayers could be investigated by analysing the protein incorporation into liposomes as described by Saliba (Saliba *et al.*, 2014).

Interaction with the HmbR receptor

The proposed role of Gly1ORF1 in iron acquisition seems to be plausible. The facts that it interacts with red blood cells and is implicated in utilising haem and haemoglobin as the sole iron source for growth of *N. meningitidis* strongly support this hypothesis. However, it is not clear how Gly1ORF1 delivers the haem-cargo into the neisserial cells. The hypothesis proposed in this study considered involvement of the haem/haemoglobin receptors HmbR or HpuAB in the transport of haem across the outer membrane. Gly1ORF1 function in this model would be limited to the role of the extracellular haem transporter delivering haem. HmbR and HpuAB receptors have been described in Section 1.2.6.3. As HpuAB receptor was not detected by PCR in strain MC58 of *Neisseria meningitidis* (which was also previously reported in the literature), the study focused on the interactions of Gly1ORF1 with HmbR receptor.

HmbR expression has been shown to correlate with invasiveness of meningococci (Harrison *et al.*, 2009; Tauseef *et al.*, 2011). Although the HmbR sequence is conserved in *Neisseria meningitidis* with sequence identity of over 90%, the phase variation of this gene results in its lack of expression in some strains. Surprisingly, Hmbr sequence homologs can be found in many Gram-negative bacteria that were also identified when searching for potential Gly1ORF1 homologs. HmbR's nearest sequence homologs outside of *Neisseria* genus with approximately 50% identical residues were identified in *Haemophilus* spp., *Mannheimia* spp., *Pasteurella* spp., *Actinobacillus* spp., *Moraxella* spp., or *Avibacterium* spp. Those Gram-negative bacteria are human or animal pathogens and in some cases can cause septicaemia.

The *hmbR* gene was successfully expressed in *E. coli*. Initial experiments showed some level of interaction of Gly1ORF1 with the *E. coli* cells producing HmbR compared to cells with an empty expression vector. This preliminary data are promising. However, more experiments must be performed to authenticate this finding. The most urgent experiments would involve pull-down experiments with the cell lysates of *E. coli* expressing *hmbR* with Gly1ORF1. It would be also interesting to invesitigate if Gly1ORF1 binding to HmbR is increased in the presence of haemin. It is possible that no difference would be observed. The

haemophore HasA in *Serratia marcenscens* binds to its outer membrane receptor HasR with similar affinity ($K_d = 5$ nM) irrespective of its haem-loaded status (Létoffé *et al.*, 2001). These experiments could not be performed in this study due to the time restrictions. However, the work on the Gly1ORF1 – HmbR interaction will be continued. If the experiments confirm the initial findings, further work will involve purification of HmbR receptor and co-crystallisation trials with Gly1ORF1 with and without haem. Further work would also involve binding affinity studies of HmbR-Gly1ORF1 interaction with and without haemin.

It is very unlikely that Gly1ORF1 and HmbR receptor could act as a typical bipartite haem/haemoglobin TonB-dependent receptor as the genes are distantly localised in the neisserial genome. Many of the typical bipartite receptors are co-transcribed from one operon and their expression is regulated by iron levels as they often contain the Fur box. The examples of such operons are HasAR form *Yersinia pestis*, HmuYR in *Porphyromonas gingivalis* or HpuAB in *Neisseria meningitidis*. However, one of the identified neisserial Gly1ORF1 homologs in *Mannheimia haemolytica* – *mha2262,* is localised close to the *hmbR2* gene (predicted to be a haem outer membrane receptor) and other genes which encode protein belonging to the haem ABC-transporter family (Fig. 5.3). The characterisation of this homolog and its interaction with the HmbR receptor in *Mannheimia haemolytica* is a part of another project in the collaboration with the DEFRA.

HmbR expression is upregulated in the iron-replete condition as are many other iron receptors in *Neisseria*. It contains the Fur box in the promoter region. However, the iron-independent mechanism of *hmbR* gene-regulation based on the MisS/R two-component system activates the gene expression as a response to uncharacterised environment stimuli (Zhao *et al.*, 2010). Gly1ORF1 does not possess a Fur box in the promoter region and there are no reports of Gly1ORF1 up-regulation in response to iron limitation. However, a 150 bp long sequence referred to as the Contact Regulatory Element of *Neisseria* (CREN) was identified upstream of the *gly1ORF1* gene. The CREN is believed to be involved in the regulation of gene expression in response to the adhesion to the host cells. The study by Deghmane and co-workers demonstrated that *gly1ORF1* expression is indeed upregulated in the cell associated bacteria

(Deghmane *et al.*, 2003). Therefore, neisserial *gly1ORF1* expression could be upregulated upon bacterial contact with red blood cells. It has been reported that *N. meningitidis* binds to RBCs utilising the erythrocyte complement receptor 1 (Brekke *et al.*, 2011). The independence of Fur up-regulation of *hmbR* could be also an adaptation to the iron acquisition in response to *Neisseria* binding to red blood cells. This could explain the higher invasiveness of the strains that express *hmbR*.

It is also possible that Gly1ORF1 might play a role of storage protein for haem. Gly1ORF1 has been shown to form dimers and higher molecular weight structures and was detected in the outer membrane vesicles. It is possible that Gly1ORF1 with bound haem is attached to the surface of the cells and acts as a reservoir of haem. It is also suspected that Gly1ORF1 might undergo a domain swapping oligomerization process as discussed above. It cannot be ruled out that the monomeric and dimeric form of Gly1ORF1 can play different roles during meningococcal pathogenesis. A similar observation was seen in the case of HasA haemophore from *Serratia marcenscens* as previously mentioned in Section 1.2.6.2. The protein as a monomer binds haem with high affinity and interacts with the outer membrane receptor HasR. The HasA can be also present as a dimer (DHasA) formed by the domain swap. The DHasA is not able to interact to HasR but was found attached on the surface of the cells (Czjzek *et al.*, 2007). Therefore, it would be interesting to investigate if Gly1ORF1 dimers and monomers differ in the affinity to haem and HmbR receptor. It is also possible that as a monomer, Gly1ORF1 might be responsible for the interaction with red blood cells and EPB4.2, whereas upon binding of haem it would form a dimer and function as a haem transporter.

Immunogenic studies on Gly1ORF1

Septicaemia and bacterial meningitis caused by *Neisseria meningitidis* are still a serious concern for parents and health professionals despite available antibiotic therapies. So far, prevention by vaccination has proven to be the most effective way of controlling the invasive meningococcal disease (IMV) especially in infants and toddlers with the highest rate of IMV incidence.

The MenAfriVac conjugate vaccine, which was designed specifically for the serogroup A strain causing the epidemics in the so called 'Meningitis Belt' in Africa, not only protects vaccinated individuals from carriage and disease, but also significantly reduced the carriage rate in unvaccinated individuals (Frasch *et al.*, 2012; Kristiansen *et al.*, 2013). The introduction of a routine vaccination program with the MenC vaccine against serogroup C eradicated the disease caused by this serogroup in England, Wales, Australia and Ireland (Andrews and Pollard, 2014). Serogroup B, therefore, became the main causative agent of meningococcal disease in high-income countries.

The first available MenB vaccines were based on outer membrane vesicles, which are produced by some bacteria during growth. The immunogenicity of such vaccines is conferred by the presence of the endogenous neisserial surface antigens, mainly PorA. The tailor-made vaccines that were developed during meningococcal epidemics in Norway (MenBVac) (Bjune *et al.*, 1991), Cuba (Va-Mengococ BC) (Sotolongo *et al.*, 2007) or New Zealand (MenZB) (Oster *et al.*, 2005) proved to be very effective in controlling the local outbreaks. However, the main drawback of these vaccines is the narrow strain specificity, limited by the presence of a monovalent PorA antigen which is extremely heterogeneous between strains (Tan *et al.*, 2010). Also, the vaccine based on PorA, which undergoes phase variation, poses a risk of antigenic selection for strains that are not covered by the vaccine.

An innovative approach based on the bioinformatics analysis of the MC58 genome resulted in the development of a new universal vaccine 4CMenB, commercially known as 4CMenB or Bexsero. The reverse vaccinology approach identified initially over 600 genes encoding extracellular antigens. Of these over 350 were expressed in *E. coli* and used to vaccinate mice to select for antigens with the highest immunogenicity and ability to elicit bactericidal response. Five antigens were selected for the vaccine formulation: factor H binding protein in complex with GNA2091, neisserial adhesion A protein (NadA) and neisserial heparin binding protein (NHBA) in fusion with GNA1030. In addition, the 4CMenB vaccine contains outer membrane vesicles that were used in the MenZB vaccine with PorA protein serosubtype B (Serruto *et al.*, 2012)

As previously mentioned, the strain coverage of Bexsero was predicted to be less than 80%, which poses a risk of positive selection for strains that are not covered by the vaccine (Vogel *et al.*, 2013). The emergence of invasive strains that do not express the main vaccine antigens like factor H binding protein as reported by Lucidarme and co-workers (Lucidarme *et al.*, 2011) is a cause for concern with regards to the efficacy of the vaccine in the future. Therefore, the search for new antigens with higher strain coverage continues.

For successful vaccine formulation the candidates should meet a number of criteria which include: ability to elicit immunogenic response in humans, reaching protective bactericidal titres and cell surface expression allowing for efficient activation of complement mediated killing. To ensure broad strain coverage, the antigens must ideally be expressed in all strains with a high degree of sequence identity. From a safety point of view, antigens showing cross-reactive activity with a host protein should be avoided due to the risk of an auto-immune response. Finally, after meeting all those criteria, the candidate should be suitable for large-scale production (Sathyamurthy, 2011).

The experiments performed in our laboratory (Meadows, 2004) and bioinformatic analysis of the available neisserial genomes proved that Gly1ORF1 is highly conserved and widespread in neisserial genomes guaranteeing broad strain coverage. The Gly1ORF1 has no homologs in the human genome. This indicates that it is extremely unlikely that the protein would elicit cross-reactive immune response with human proteins.

The data in the pilot study show that Gly1ORF1 can elicit an immunogenic response in humans, as high antibody titres against Gly1ORF1 were detected in the sera one patient with invasive meningococcal infection. It would be interesting to test larger numbers of sera, however samples from septic patients are extremely rare. Also, a high response to Gly1ORF1 was detected in sera from patients with no suspected meningococcal infection. The fact that not all studied sera from septic patients showed a high antibody response to Gly1ORF1 can be attributed to the fact that the direct ELISA method used in this study has major drawbacks of a high background in negative samples and low specificity. A more sophisticated method for the detection of antibodies

against factor H binding protein - fHBP (a component of Bexsero vaccine) was described by Ala'Aldeen and co-workers (Ala'aldeen *et al.*, 2010). The Luminex-based immunoassay used in this study allowed very sensitive quantification of anti-fHBP antibodies in sera of septic patients greatly reducing the background. It detected subtle differences in antibody levels even in highly diluted sera, which is not possible with the direct ELISA method. Also, this method allowed reliable monitoring of the changes in the antibody levels against the antigen of interest throughout the course of infection. Ala'Aldeen and co-workers showed that antibody levels against fHBP varied significantly between septic patients. In most of the cases of invasive meningococcal disease the levels of anti-fHBP antibodies were rising after admission to the hospital; however, in two individuals there was no increase in response to the antigen during infection and convalescence, although the antigen was expressed in both strains responsible for the disease. Also, it has been shown that antibody levels, after reaching a peak during infection, were decreasing during long term convalescence except in one case. So, the fact that in this study not all isolates showed a high response to Gly1ORF1 with no obvious increase in the antibody titres after a month from the onset of symptoms is not unusual and can be explained by the differences in the immunogenic response between affected individuals.

The ELISA test shows a response to Gly1ORF1 and OMV in most of the samples from patients with no suspected meningococcal disease. It is not surprising, as asymptomatic carriage has an immunogenic effect (Jordens *et al.*, 2004; Yazdankhah and Caugant, 2004) and most people at some point in their lives were colonised by *N. meninigitidis.* The fact that the mean response to OMV was higher than to Gly1ORF1 in the tested samples is something that was expected as OMVs contain many antigens and are known to be strongly immunogenic. However, the higher response to Gly1ORF1 than to OMVs in some individuals can indicate that Gly1ORF1 can also be highly immunogenic and elicit an antibody response as a result of carriage. There is also a chance that the individuals showing a higher response to Gly1ORF1 than to OMVs were colonised with a strain that was very divergent from MC58 *N. meningitidis* in the OMV antigen composition and sequence. Thus, detected antibody

response to MC58 OMVs could be lower than that of the colonising strain, whereas the response to a conserved Gly1ORF1 would be the same irrespective of the colonising strain. It would also be very interesting to see if the high Gly1ORF1 antibody levels in those samples correspond with high bactericidal titres of the sera against *N. meningitidis*. However, the low volume of the samples obtained from the Health Protection Agency limited the possibility of investigating it further.

Previous initial experiments in our laboratory to test if Gly1ORF1 antibodies have a bactericidal effect on *N. meningitis* gave promising results (Sathyamurthy S., 2011). This subject was therefore investigated in more detail. The data in this study shows that mouse anti-sera raised against soluble Gly1ORF1 meet current Health Protection Agency criteria to be regarded as protective against meningococcal infection as they reach 50% killing in a serum dilution of up to 1:16. The next step would also involve testing if the bactericidal effect of the Gly1ORF1 can provide broad strain protection. In this study only two strains were compared for susceptibility to anti-Gly1ORF1 antibodies: MC58 and H44/76 and showed no statistically significant difference.

Most crucial for the activation of complement-mediated killing is the availability of the antigen to the bactericidal antibodies circulating in the blood. This study showed that the levels of expression of Gly1ORF1 are probably very low as the expression of the protein was not readily detectable with polyclonal antibodies. Therefore, it is still important to generate monoclonal antibodies against Gly1ORF1 for future experiments. Although Gly1ORF1 expression levels might be low, it does not mean this protein should be disregarded as a good immunogen. The expression of other surface antigens, like factor H binding protein (hHBP) can vary significantly between strains (McNeil *et al*., 2009). However, according to the literature, there is no correlation between levels of surface expressed hHBP and the antibody response elicited against this protein during invasive infection (Ala'aldeen *et al*., 2010). Also, the fact that tested strains MC58 and H44/46 were much more susceptible to complement mediated killing than the MC58 *gly*1ORF1 *null* mutant strain indicates that the levels of protein were sufficient for eliciting a bactericidal response in serum.

Finally, Gly1ORF1 without the histidine-tag can be obtained with high yields as inclusion bodies with a reduced number of purification steps. This might be significant for large scale production of the antigen for vaccine formulation.

The data show some promising features of Gly1ORF1 to be considered for further studies evaluating its potential as a vaccine candidate.

Alpha protein – another promising vaccine candidate?

Just by chance this study identified that another neisserial protein may have potential for vaccine development: the alpha protein of IgA1 protease. This protein was subject to an earlier study in our lab (Parsons, 2003) and was shown to be expressed in the majority of clinical isolates tested. Thus, antibodies against this protein were used as a positive control for FACS analysis when seeking to measure surface expression of Gly1ORF1 on MC58.
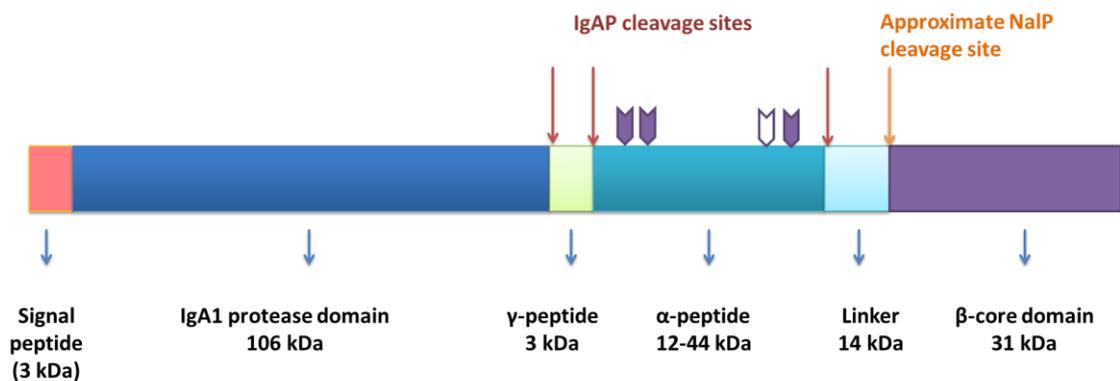
Figure 6.3 Schematic representation of the product of *iga* gene.

The arrows show different autoproteolytic cleavage sites (red arrows) and NalP cleavage site (orange arrow). The purple arrows represent Nuclear Localisation Signals (NLS) in long alpha-proteins containing four NLS, where the open arrow indicates NLS in short alpha-peptides with only one NLS. The proteolytic cleavage generates mature IgA1 protease and different variants of alpha protein depending on the site of cleavage.

Alpha protein is expressed as a part of the IgA1 protease precursor protein, which undergoes auto-proteolytic cleavage to yield the mature IgA1 protease, β-

core translocator domain and variable alpha-protein (Fig. 6.3). Recent studies also showed involvement of NalP proteases in the variable processing of the *iga* gene. The NalP cleavage yields different variants of alpha protein which are either released outside the cells, where they remain covalently attached to the translocator domain or loosely attached to the outer membrane (Roussel-Jazédé *et al.*, 2014; Roussel-Jazédé *et al.*, 2013). The role of IgA1 protease in neisserial pathogenesis has been studied extensively (Ayala *et al.*, 2001; Beck and Meyer, 2000; Hauck and Meyer, 1997; Lorenzen *et al.*, 1999). Alpha protein function, in turn, remains highly enigmatic. Although the alpha protein is highly variable in length, it contains a varying number of 10-amino acid long repeats which are highly conserved between strains of *Neisseria* spp. Moreover, the crystal structure of the neisserial alpha-protein solved in our laboratory (Parsons, 2003) showed that it forms amphipathic alpha helices. The structure of alpha protein has many features similar to the motifs that bind to MHCII class molecules of T-cells (Jose *et al.*, 2000), which indicates that it might have strong immunogenic properties. This was confirmed by the initial results in this study were the anti-alpha antibody titers in the tested sera were comparable or even higher than the response to outer membrane vesicles. Our initial data also demonstrate that the anti-alpha antibodies might have a strong bactericidal effect on *N. meningitidis,* however this needs to be further confirmed by comparison of pre- and post-immunization sera which were not available at the time of the study.

Our experiments were a preliminary investigation into the potential of alpha protein in vaccine development. In-depth analysis of alpha protein was set up as a separate project in our laboratory. It will be also investigated if the fusion of Gly1ORF1 and alpha protein increases the serum bactericidal titre. The *gly1ORF1* gene is present in all analysed clinical isolates with over 90% identity, whereas alpha protein has an advantage of high immunogenicity (Jose *et al.*, 2000). In addition, alpha protein is also highly conserved in neisserial species and can be detected in most of the clinical isolates tested in our laboratory. Indeed, IgA1 protease expression levels have been correlated with virulence in *N. meningitidis* (Vitovski *et al.*, 1999). Such fusion of two antigens was used in the recently approved Bexsero (4CMenB) vaccine. In the

formulation of the vaccine factor H binding protein, fusion with GNA2095 and neisserial heparin binding antigen and fusion with GNA1030 were used to stabilise the antigens and increase their immunogenicity (Serruto *et al.*, 2012). It can be hypothesised that the fusion Gly1ORF1 and alpha protein could induce production of protective antibodies that would give broad coverage of strains and could be a basis for the development of universal meningococcal vaccine against multiple serotypes (Patent WO 2013079970 A1).

Identification of Gly1ORF1 homologs

The Gly1ORF1 protein sequence was originally thought to be unique to the *Neisseria* when the protein was first described in 1999 (Arvidson *et al.*, 1999); however, as the databases have expanded, subsequent searches for structural homologs reveal several interesting protein sequences that are encoded in the genomes of other human and animal pathogens. This study focused on characterization of Gly1ORF1 homologs and paralogs found in *Haemophilus influenzae*, *Mannheimia haemolytica* and *Neisseria meningitidis*.

Since 2012 a number of new potential homologs were added to the databases. Many of these proteins are encoded by genomes of human pathogens with emerging clinical importance due to the increasing incidence of bacteraemia and drug resistance. The most interesting examples are: *Proteus mirabilis* and *Providencia stuartii,* which are human uropathogens (Armbruster *et al.*, 2014); *Acinetobacter baumanii* (Doughari *et al.*, 2011) or *Morganella morganii* (Lin *et al.*, 2014). Some of the pathogens are associated with hospital-acquired infections and target immunocompromised patients (Williams *et al.*, 1983). The closest Gly1ORF1 homolog, with 51% sequence identity, was recently identified in a novel species of *Haemophilus* – haemolytic *H. sputorum* (Nørskov-Lauritsen *et al.*, 2012).

The similarity of the predicted secondary structure of these proteins, and the fact that they were found in pathogens causing septicaemia, suggests that together with neisserial Gly1 they might constitute a novel group of virulence factors probably being a result of convergent evolution of pathogens colonising similar niches. The fact that the *mha2262* gene in *Mannheimia haemolytica* (referred to as Mannhemia Gly1ORF1 paralog) is also in close proximity to a

gene involved in haem transport suggests that this group of proteins could be also responsible for iron acquisition.

Our attempts to solve the crystal structure of Gly1ORF1 homologs in *M. haemolytica* and *H. influenzae* and neisserial Gly1ORF1 paralog were unsuccessful. It has been shown in this study that homologs and the Gly1ORF1 paralog are able to form homo-oligomeric structures.

The structure of Gly1ORF1 paralog in *M. haemolytica (*MHA2262*)* was solved as a part of the parallel project in the laboratories of Profs. Artymiuk and Sayers. The Gly1ORF1 paralog structure showed a similar fold to neisserial Gly1ORF1 and also forms a dimer. However, there was no observed haemin binding by this paralog (unpublished data, verbal communication from H. McMellon). This is quite surprising as the gene encoding this paralog lies next to the gene encoding an HmbR2 receptor-like protein. On the other hand, neisserial Gly1ORF1 paralog did not show convincing evidence of haemin binding either. Therefore, it is conceivable that the Gly1ORF1 paralogs have different functions to Gly1ORF1 and its homologs despite some structural similarity.

This hypothesis is supported by the recently published data on the neisserial Gly1ORF1 paralog, also referred to as Adhesin Complex Protein (ACP) (Hung *et al.*, 2013). The study showed that ACP is involved in adhesion; the ΔACP mutant strain of *N. meningitidis* adhered to different types of mammalian cells to a lesser extent than the WT (up to 75% reduction). The protein has been shown to be surface exposed (Hung *et al.*, 2013) and was also identified in outer membrane vesicles (Ferrari *et al.*, 2006; Post *et al.*, 2005). Although ACP was reported to be up-regulated in iron-replete conditions (Grifantini *et al.*, 2003), no obvious motif corresponding to the Fur box could be detected upstream of the sequence. Similarly to Gly1ORF1, ACP is widely distributed in *Neisseria.* It is highly conserved and the ACP variants between strains differ only with 1 or 2 residues. ACP was shown to be highly immunogenic and the ACP-immunised mouse sera were eliciting high bactericidal titres (up to 1024) (Hung *et al.*, 2013). It would be interesting to see if the paralog in *M. haemolytica* (also predicted to be adhesin) shares similar functions and if it can also elicit strong

immunogenic responses. This is the objective of another project in the Sayers' laboratory; investigating potential vaccine candidates in *Mannheimia haemolytica*. This would also be beneficial for identification of potential vaccine targets against other pathogenic bacteria that encode members of the Gly1ORF1-paralog protein family. An ACP sequence homolog can be found by PSI-Blast search in *Moraxella catarrhalis*, which is an emerging human pathogen. It is the third most common cause of the acute otitis media in children and it becomes much more prevalent after introduction of anti-pneumoccocal vaccine (Bernhard *et al.*, 2012). This pathogen is of rising clinical importance and imposes an increasing economic burden; however, no vaccine has yet been developed against it.

Gly1 homologs do not share high sequence homology. However, there are numbers of residues that are conserved in the sequences (Fig. 5.1 and 5.2). As the structures of the homologs could not be solved, the structure of neisserial Gly1ORF1 K12A mutant was used as a model to identify the position of conserved residues.

Most of the residues with over 50% conservation are localised between the two β-sheet domains of the monomer. The residues are localised in the Gly1ORF1 structure in the cavity that has not been identified as a potential haem-binding site in the docking experiments.

Many of these residues were subjected to point mutations in this study. However, the replacement of these residues did not affect the ability of neisserial Gly1ORF1 to bind haem. This might indicate that the main function of the Gly1ORF1 homolog family is not haem binding in *Neisseria*. This conserved site might be responsible for interactions with other host factors, e.g. EPB4.2. However, the FACS analysis did not show convincing results for the interactions of the HinfGly1ORF1 homolog and neisserial paralog with red blood cells. It does not rule out the possibility that the protein interacts with the EPB4.2. The proteins might be transported into the red blood cells by the outer membrane vesicles. HinfGly1ORF1 and ManhGly1ORF1 subcellular localisation prediction could also be indicative of OMV association, but it is something that would need to be confirmed experimentally.

Gly1ORF1 homolog protein family is a very interesting, uncharacterised group of proteins that may be involved in pathogenesis. However, more work needs to be done to confirm the structural similarities between these proteins and to determine their function in the interaction with the host. The facts that many of these homologs are widely spread and conserved within the species encoding them and are surface exposed and immunogenic raises a possibility of their potential as vaccine targets. It is something that is definitely worth further investigation.

CHAPTER 7   Materials and Methods

7.1   Bacterial strains, vectors and culturing conditions

Escherichia coli M72 (*Δbio-uvrB*) *lacZ(Am) rpsL Δ(trpEA2) (λ Nam7 Nam53 cI857 ΔH1*) (O'Connor and Timmis, 1987) was used to express Gly1 constructs cloned into the pJONEX4 plasmid (Sayers *et al.*, 1992) which had been modified by inserting a hexa-histidine-encoding cassette to generate histidine-tagged proteins (Meadows, 2004). Genes cloned into the pJONEX4 vector are under the control of the bacteriophage promoter λP_L which is repressed at lower temperatures by the chromosomally encoded repressor cI857. Therefore the *E. coli* M72 cells were grown at temperatures below 30°C in Luria-Bertani (LB) media containing ampicillin at a final concentration of 100 µg/ml. Induction of protein expression was achieved by increasing the culture temperature to 42°C, which results in the dissociation of the repressor from the promoter region allowing transcription to occur.

E. coli XL1 Blue (*endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[:: Tn10 Tet^R proAB$^+$ lacI^{q}Δ(lacZ)M15] hsdR17(r$_K$- m$_K$$^+$)*) was used to propagate the other vectors used in this study and M13mp18 bacteriophage (Sayers *et al.*, 1992). This strain is characterised by absence of endonuclease (*endA*) and recombination activity (*recA*) resulting in higher yields of cloned DNA and stable inserts. Mutation in the gene *hsdR* additionally prevents cleavage of the cloned DNA by the *EcoK* endonuclease system. Presence of the F' plasmid also enables infection M13mp18 bacteriophage. *E. coli* XL1 Blue cells were grown at 37°C in 2xYT in the presence of tetracycline at a final concentration of 12.5 µg/ml with the addition of suitable cognate antibiotics (e.g. ampicillin) when a vector was present within the cells.

E. coli BL21(DE3) *(fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5)* and *E.coli* BL21 Star^{TM} (DE3) (*ompT hsd S$_B$ (r$_B$-m$_B$-) gal dcm rne 131 (DE3)*) (Invitrogen Life Sciences) were used for high yield expression of proteins from the vector pET21a+ (Novagen). Genes cloned in the pET21a+ vector are under control of strong promoter derived from T7 bacteriophage. Expression of a

cloned gene can be repressed by addition of glucose to the growth media at the concentration of 0.1 % (w/v). Induction of gene expression was initiated by addition of isopropyl-beta-D-thiogalactopyranoside (IPTG) to the final concentration of 0.1 mM.. *E. coli* BL21 Star™ (DE3) strain was only used when expression from BL21(DE3) was undetectable or low; *E. coli* BL21 Star™ (DE3) has a truncated version of RNaseE which is unable to degrade the mRNA template and thus shows increased stability of mRNA for heterologous genes cloned into the expression vector. This can result in improved yields of recombinant proteins.

7.2 Preparation of $CaCl_2$ competent cells

The desired strain of *E. coli* was streaked on to a LB agar plate supplemented with the appropriate antibiotics (if required); this was then incubated overnight at an appropriate temperature for the *E. coli* strain being used.

A single colony was picked up from the agar plate inoculated into 3 ml of LB media supplemented with the appropriate antibiotic (if required) and grown overnight at the appropriate temperature.

The following day 100 ml of LB media (with antibiotic if required) was seeded with 1 ml of the overnight culture and grown until the A_{600nm} reached approximately 0.4.

The culture was centrifuged at 3000 x g for 10 min at 4°C and the supernatant was discarded. Pelleted cells were resuspended in 30 ml of ice cold sterile 0.1 M $MgCl_2$ and then centrifuged again at 1000 x g for 20 min at 4°C.

The supernatant was discarded and cells were gently resuspended in 5 ml of sterile ice cold 0.1 M $CaCl_2$ by gentle pipetting. Once resuspended, another 35 ml of sterile ice cold 0.1. M $CaCl_2$ was added. Cells were incubated for 1 hour on ice and then pelleted by centrifugation at 1000 x g for 20 min at 4°C. The competent cells were gently resuspended into 6 ml ice-cold sterile 85 mM $CaCl_2$ containing 20% (v/v) glycerol and aliquoted into pre-chilled tubes. Competent cells were stored at -80°C.

7.3 Cloning

7.3.1 Plasmid DNA preparation

Plasmid DNA was prepared using a DNA purification kit E.Z.N.A. (Omega Bio-Tek). The protocol was provided by the manufacturer.

7.3.2 PCR

The PCR mix composition used for all genes amplification in this study was containing: 25 µl of 2xPCR mix (Thermo Scientific), 2 µl of template DNA (approximately 50 ng), 2 µl of each primer at 10 µM concentration and deionised water up to 50 µl.

Oligonucleotides used in this study are summarised in the Tables 7.1-7.3 with the oligonucleotide's sequences, restriction enzyme recognition sites and PCR conditions.

Table 7.1 Oligonucleotides sequences for amplification of haemoglobin receptor genes.

Primer sequence ID	Primer sequence	Restriction enzyme – recognition site
HmbR_for	d(TTATACATATGAAACCATTACAAATGCTCCCTA)	NdeI
HmbR_rev	d(TATTAAGCTTTCAAAACTTCCATTCCAGCG)	HindIII
HpuA_for	d(ATGCCATATGAAATACAAAGCCC	NdeI
HpuA_rev	d(TTAGAAGCTTCCTTATTTTGTGAGATGGG)	HindIII
HpuB_for	d(TAATCATATGCCCATCCCCTTCA)	NdeI
HpuB_rev	d(TATTAAGCTTGCTTAGAACTTCGCTT)	HindIII

*PCR program: initial denaturation at 95°C for 1 minute, 25 cycles of denaturation at 94°C for 1 minute, annealing at 55°C for 1 minute and elongation at 72°C for 3 min 30 seconds finished by final elongation for 3 minutes.

Table 7.2 Oligonucleotides sequences for amplification of Gly1ORF1 homologs and neisserial Gly1ORF1 paralog genes.

Name of the oligonucleotide	Sequence of the Oligonucleotide	Restriction enzyme recognition sites
Gly1_man.f	d(TTTCGAATTCTAGAGGAAACAAAAATGCGTAAATTATTAGTAATT)	EcoRI
Gly1_man.h6.r	d(TGTTGGATCCCTAAAGTATTCATCAAATGAACATC)	BamHI
Gly1_man.r	d(ACAAAAGCTTACTTAAAGTATTCATCAAATGAACATC)	HindIII
Gly1_ManS2.f: EcoR1	d(TTTAGAATTCTAAGGAGTTACATTTATGAGTAAATTATTAGTAATTACTGC)	EcoRI
Hinf_gly1.r	d(TATTCGATCCCTCAAGCTTTCATACTGCGACGT)	HindIII
Hinf_Gly1H6.r	d(AAATGTCGACGCTTATTTTATCTCTATTTACGCTTATGA)	SalI
Gly1_hinf_K.f	d(AATAGGTACCTGAGGAGAAACAAATGACAAAATTACTCACTCACTATTGGAGC)	KpnI
NGly_Par_for	d(AAGGAAAACATATGAACTTCTGACCACC)	NdeI
NGly_Par_rev	d(ATTTAAGCTTCATTAACGTGGGGAACAGTC)	HindIII

* PCR program: initial denaturation:95°C for 1 min; 25 cycles of: denaturation at 94°C for 1 min, annealing for 1min at 50°C for *H. influenzae* and 55°C for *M. haemolytica*, 58°C for *N. meningitidis*, elongation at 72°C for 1 min and a final elongation at 72°C for 5 min.

Table 7.3 Oligonucleotides sequences for amplification of the native non-histidine tagged Gly1ORF1 gene.

Name of the oligonucleotide	Sequence of the Oligonucleotide	Restriction enzyme recognition sites
Gly1_NheI	d(TTTTATGGCTAGCATGTTCCTTTCTGCCGTATTG)	NheI
universal M13	d(CAGGAAACAGCTATGAC)	N/A

*PCR program: initial denaturation: 95°C for 2 minutes; 25 cycles of: denaturation at 94°C for 30 seconds, annealing at 56°C for 30 seconds, elongation at 72°C for 45 seconds and final elongation at 72°C for 2 minutes.

7.3.3 Restriction digest

Plasmid DNA and PCR products were digested by restriction enzymes (restriction enzymes were purchased from Promega) according to manufacturer's protocol. Digested DNA was separated on 1% (w/v) agarose gel in TAE buffer at 60V and bands corresponding to digested vectors and inserts were excised from the gel. The DNA in the excised agarose gel was then purified using the QIAEX II Gel Extraction Kit (QIAGEN) according to the manufacturer's instructions.

7.3.4 Ligation

Band purified DNA was ligated at a molar ratio of 3:1 (insert : vector) with 2.5 U of T4 ligase (Thermo Scientific) in a final volume of 20 µl in the buffer provided by the manufacturer. Ligation reaction was incubated for 2 hours at 25°C followed by heat inactivation at 70°C for 5 minutes.

7.3.5 Transformation

An aliquot of 10 µl of ligation product was added to 100 µl Ca-competent *E. coli* M72 cells. Cells were incubated on ice for 1 hour with occasional gentle mixing and then plated onto LB agar plate supplemented with the appropriate antibiotics. Plates were incubated overnight at 30°C.

Single colonies were used to inoculate 5 ml of LB media with the appropriate antibiotics and these cultures were grown overnight at the appropriate temperature. Plasmid DNA was isolated from the centrifuged cells using the E.Z.N.A Mini Kit II (Omega Bio-Tek) according to the manufacturer's protocol.

Presence of the insert was confirmed by sequencing of the plasmids at the Core Genomic Facility, School of Medicine, Sheffield.

7.4 Site directed mutagenesis

7.4.1 Site directed mutagenesis by phosphorothioate approach

A bacteriophage M13mp18 derivative carrying the wild type *glyORF1* gene was used as the template for site directed mutagenesis; this was kindly provided by Sathyamurthy. The method described previously (Sayers *et al.*, 1992) was followed, with minor modifications.

7.4.1.1 Propagation of M13mp18 in XL1Blue cell cultures and preparation of single stranded DNA template

Serial ten-fold dilution of bacteriophage were prepared in sterile distilled water and 5 µl of $1:10^7$ to $1:10^9$ dilutions were added to 200 µl of log phase *E. coli* XL1 Blue cell culture grown in 2YT media supplemented with 12.5 µg/ml tetracycline. The *E. coli* XL1 Blue cells and bacteriophage were mixed with top agar and

poured onto LB plates supplemented with tetracycline at the final concentration of 12.5 µg/ml. Plates were then incubated overnight at 37°C.

The following day a single plaque was transferred into 100 µl of fresh log phase XL1 Blue culture in 2YT with tetracycline (12.5 µg/ml) and incubated overnight at 37°C. Large scale bacteriophage preparation was prepared by adding the 100 µl of overnight propagated bacteriophage into 100 ml of fresh culture of XL1 Blue at the optical density A_{600nm} of 0.3 and incubation at 37°C for 5 hours.

Bacteriophage was separated from the *E. coli* cells by centrifugation at 3500 x *g* for 20 min at 4°C. The bacteriophage was then precipitated from the resulting supernatant by adding 1/5 volume of 20% PEG 6000 (w/v) in 2.5 M NaCl and incubating at 4°C overnight. Precipitated bacteriophage was pelleted by centrifugation at 3500 x *g* for 20 min at 4°C. The pellet was then resuspended in 1 ml of TE buffer (10mM Tris-HCl pH 8, 1 mM EDTA) and 9 ml of sterile distilled water and then centrifuged again at 3500 x *g* for 20 min at 4°C. The supernatant was transferred into a clean tube and bacteriophage was precipitated again by adding 2.2 ml of 20% PEG 6000 (w/v) in 2.5 M NaCl and incubating at 4°C for another 30 min followed by centrifugation at 3500 x *g* for 20 min at 4°C. The pellet was then resuspended in 500 µl of NTE buffer (100 mM NaCl; 10 mM Tris-HCl, pH8; 1 mM EDTA). DNA purification from bacteriophage proteins was performed using standard phenol-chloroform extraction and the resulting DNA was precipitated with standard ethanol precipitation. The precipitated DNA pellet was resuspended in 50 µl of TE buffer (10 mM Tis-HCl, 1 mM EDTA) and concentration of DNA was determined using UV spectrophotometer; 37 µg/ml of single stranded DNA gives an absorbance of 1AU at 260 nm.

7.4.1.2 Homoduplex of mutant DNA preparation

Phosphorylation of primers

Custom synthesized oligonucleotides were resuspended in TE buffer and diluted to 5AU at 260nm. A mixture of 2 µl primer, 6 µl 500 mM Tris-HCl, pH 8, 3 µl 100 mM $MgCl_2$, 2 µl 100 mM DTT, 3 µl ATP, 5 U of T4 Polynucleotide Kinase and 14 µl of sterile distilled water was incubated for 15 min at 37°C and heat then inactivated at 70°C for 10 min. The sequences of the oligonucleotides

are shown in Table 7.4. Nucleotides encoding introduced mutant codons are underlined.

Table 7.4 Oligonucleotides used for SDM by phosphorothioate method.

Primer ID	Primer sequence
Gly_G47R	d(CCTTGCCACCTGTCGGAACGCCGCAACAGGTCAGCTTTGCT)
Gly_R48G	d(CCTTGCCACCTGTCGGACTCCCCCAACAGGTCAGC)
Gly_K12E	d(GATTTTTTGGACTTCTATGTA<u>TTC</u>GTTGTTGTCCGTTTTACA)
Gly_Y26F	d(CCGCCATGCCGAACGA<u>AAA</u>TTCGTAAAGATTGCGGTT)
Gly_F28R	d(CTTTTTTTGCCGCACTGCC<u>GCG</u>CGAATATTCGTAAAG)
Gly_E35K	d(GCTGTTGCGTATGGCAAT<u>TTT</u>TTTTTTTTGCCGCACT)
Gly_N67D/A	d(GGTGTACATAAATTCGCC<u>GKC</u>TTGGAATTTCATCGTTGC
Gly_Y72F	d(CGAAGCCTGTCCATATGGT<u>GAA</u>CATAAATTCGCCGTTTTG)

Annealing of primers

A mixture of 10 µl 500 mM Tris-HCl, pH 8; 10 µl 500 mM NaCl, 3 µl the phosphorylated primer (as described above) and 10 µg of single stranded M13mp18Gly1ORF1 DNA was prepared and the final volume was adjusted to 36 µl with sterile distilled water. The mixture was incubated in a water bath for 5 min at 80°C and then allowed to cool slowly to room temperature over a period of at least an hour.

Polymerisation

The following were added to the tube containing the phosphorylated primers (see above); 10 µl 10 mM ATP, 10 µl 100 mM MgCl$_2$, 3 µl 500 mM Tris H-Cl, pH 8; 5 µl 1 M DTT; 2.5 µl 10 mM dATP, 2.5 µl 10 mM dGTP, 2.5 µl 10 mM dTTP, 2.5 µl 10 mM dCTPαS, 10 U of Klenow fragment, 15 U of T4 DNA ligase (Thermo Scientific) and sterile distilled water to a final volume of 100 µl. Polymerisation was carried out overnight at room temperature.

Treatment with T5FEN Exonuclease

2 µg of T5FEN exonuclease were added to the polymerisation reaction to remove linear and nicked DNA; the mixture was incubated at 37°C for 30 min and the reaction was stopped by incubation at 70°C for 10 min.

Nicking reaction

A mixture containing 10 µl of T5 endonuclease digested DNA (from the previous step), 5 µl of 100 mM DTT, 5 µl of 500 mM NaCl, 1 µl of 500 mM Tris-HCl, pH 8; 4 µl of 100 mM $MgCl_2$, and 12.5 U AvaI restriction enzyme was prepared and the final volume was adjusted to 50 µl with sterile distilled water. The reaction was allowed to proceed for 90 min at room temperature and was then heat inactivated by incubation at 70°C for 10 min.

DNA Gapping

Gapping of the nicked DNA was achieved by incubation of 100 µl of reaction mixture from previous step with 40 U of gene 6 of T7 exonuclease at 37°C for 60 min. The reaction was then inactivated by incubation at 70°C for 10 min.

Generation of Double-Stranded DNA

Double stranded DNA was obtained by incubation of 100 µl of the gapping reaction (from above) with 10 µl of 10 mM ATP, 2.5 µl of 10 mM dNTPs (10 mM of each: dATP, dGTP, dCTP, dTTP), 10 U of T4 DNA ligase and 5 U of DNA polymerase I at 37°C for 3 hours.

7.4.1.3 Transformation of the M13mp18Gly1 into XL1 Blue cells

Double stranded DNA containing desired mutation was transformed into $CaCl_2$ competent *E. coli* XL1 Blue cells. Briefly, 10 µl of double-stranded DNA was added to 100 µl of $CaCl_2$ competent *E. coli* XL1 Blue cells and incubated on ice for 1 hour with occasional gentle mixing. Transformed cells were then mixed with previously prepared 250 µl of fresh mid-log culture of XL1 Blue (A_{600nm}=0.3) and added to 3 ml of LB top agar containing 12.5 µg/ml of tetracycline kept at 50°C. The top agar containing the transformed XL1 Blue was poured over a LB agar plate supplemented with 12.5 µg/ml tetracycline. Plates were allowed to set at room temperature and were then incubated overnight at 37°C.

7.4.1.4 Miniprep of phage DNA

To identify clones with the desired mutation small scale isolation of phage DNA was performed. Single plaques were transferred into 6 ml aliquots of fresh culture of XL1 Blue at the A_{600nm} of 0.3 and were incubated at 37°C for 5 hours.

Cells were centrifuged down at 13000 x *g* for 1 minute and double stranded DNA (RFIV) was isolated using Omega E.Z.N.A. Plasmid Miniprep kit as directed by the manufacturer's instructions. DNA was sequenced at the Core Genomic Facility at the Medical School, University of Sheffield.

7.4.2 Site directed mutagenesis by PCR

A single step polymerise chain reaction (PCR) based site directed mutagenesis was used to make mutants of Gly1ORF1; the methodology was based on the protocol for Q5 Site Directed Mutagenesis (New England Biolabs) with modifications. The pJONEX4-Gly1 C-histidine tag plasmid, previously synthesised in the Sayer's laboratory by Meadows (Meadows, 2004), was used as a template for the creation of single-amino-acid substitutions. The vector pET21-Gly1ORF1 was used as a template for the deletion of the signal peptide in the native Gly1ORF1 construct in T7-inducible vector.

7.4.2.1 PCR

The whole plasmid DNA was amplified using a pair of oligonucleotides which were designed to introduce the desired mutations into the gene of interest. Sequences of oligonucleotides are shown in Table 7.5. Nucleotides encoding introduced mutant codons are underlined.

Table 7.5 Oligonucleotides used for SDM by single step PCR method.

Primer ID	Primer sequence
Gly1_F4A_for	d(GGACAACAAC<u>GCG</u>TACATAGAAGTCCAAAAAATC)
Gly1_F4A_rev	d(GCCCACACGGTTTGGGCG)
Gly1_K12A_for	d(GGACAACAAC<u>GCG</u>TACATAGAAGTCCAAAAAATC)
Gly1_K12A_rev	d(GTTTTACAGGAAAACACCG)
Gly1_M63A_for	d(TTGGGCAACG<u>GCG</u>AAATTCCAAAACGGCG)
Gly1_M63A_rev	d(CGACCGCTGCCCATACCT)
Gly1_dSTOP_for	d(TCACCATCAT<u>TAATAA</u>TTGGCGTAATCATGG)
Gly1_dSTOP_rev	d(TGATGGTGTTTGTCATCG)
Gly1_del.for	d(GATACGGTGTTTCCTGTAAAAC)
Gly1_del.rev	d(CATATGTATATCTCCTTCTTAAAGTTAAC).

The PCR mixture was prepared by mixing 25 µl KAPA Hifi Hot Start PCR mix (Kapa Biosystems), 1.5 µl 10 µM forward primer, 1.5 µl 10 µM reverse primer, 0.1 µg template DNA and the volume adjusted to 50 µl with sterile distilled water. PCR was carried out in the thermal cycler (Eppendorf, Mastercycler Personal); the following PCR protocol was used: initial denaturation at 95°C for 3 min, followed by 25 cycles of denaturation step at 98°C for 20 sec, annealing at 59°C for 15 sec and elongation at 72°C for 1 min 30 sec finished by final elongation step at 72° for 3 minutes.

PCR products were separated on a 0.8% (w/v) agarose gel and linearised mutagenized constructs were band purified and gel extracted using the QIAEXII Gel Extraction Kit (QIAGEN).

7.4.2.2 Ligation of the linear plasmid and treatment with DpnI

Linear DNA PCR product (5µl of 125 ng/µl) was first phosphorylated with 5 U of T4 PNK (T4 polynucleotide kinase) in the buffer provided by the manufacturer (Promega); the reaction was carried out at 37°C for 30 min. The original template DNA was removed by treatment with DpnI restriction enzyme which cuts only methylated DNA, followed by circularisation of the vector with T4 ligase. This was achieved by adding 5 U of DpnI (Promega), 1.5 U of T4 ligase (Promega) and 1x ligase buffer to the phosphorylated linear DNA PCR product. Samples were incubated for 30 min at room temperature and heat inactivated by incubation at 70°C for 5 min.

The circularised construct with desired mutations was then transformed into CaCl$_2$-competent *E. coli* M72 cells as described previously (Section 7.3.5).

7.5 Protein production

7.5.1 Small scale production of proteins by heat induction

A single colony of freshly transformed *E. coli* M72 cells containing the pJONEX4 plasmid an insert of the choice were used to inoculate 3 ml LB media with 100 µg/ml of ampicillin and incubated overnight at 30°C in a shaking incubator. The following day, 5 ml of LB media with ampicillin were inoculated with 100 µl of the

overnight culture and incubated for 3 hours at 30°C. Protein production was induced by increasing the temperature to 42°C for 3 hours. After induction cultures were incubated overnight at room temperature. Samples were then analysed on SDS-polyacrylamide gel.

7.5.2 Large scale protein production in the heat inducible system

A fermenter containing 3 litres of 4YT with 100 µg/ml of carbenicillin and 300 µl of Antifoam 204 (Sigma) was inoculated with 200 ml of the overnight culture. The culture was incubated at 30°C with shaking at 450 rpm until optical density reached $A_{600nm} = 2$ and the temperature was increased to 42°C for 3 hours. After induction temperature was lowered to 20°C and the fermenter culture was incubated overnight. Protein production was confirmed by SDS PAGE comparing samples of cell lysates and supernatant before and after induction.

Cells were separated from the supernatant by centrifuging at 9000 x *g* for 45 min at 15 C and supernatant was transferred into sterile bottles and stored at 4°C. The cell pellets were stored at -80°C.

7.5.3 Large scale protein production in the IPTG inducible system

Freshly transformed *E. coli* BL21(DE3) with the pET21 vector containing an insert of the choice were grown on LB plates supplemented with 0.1% (w/v) glucose and 100 µg/ml ampicillin. A single colony was inoculated into 5 ml of LB liquid media with 0.1% (w/v) glucose and ampicillin (100 µg/ml) and grown overnight at 37°C. A 100 ml inoculum in the same liquid media was prepared by adding 1 ml of an overnight culture and incubation was continued overnight at 37°C.

A 1 litre flask with 4YT media containing 0.1% glucose and 100 µg/ml carbenicillin was inoculated with 50 ml of an overnight culture and was incubated at room temperature for 5 hours. Cultures were induced with 0.1 mM IPTG and incubated a further 2.5 hours at room temperature. Cultures were centrifuged at 9000 x *g* for 30 minutes at 4°C. Induced cell pellets were stored at -80°C.

7.5.4 Protein production in the auto-induction culture medium

Protocols for the auto-induction of genes expressed from T7 promoter expression vector were previously described by Studier (Studier, 2005).

Auto-induction medium ZYM-5052 used for this study was prepared by mixing 959 ml of autoclaved ZY medium (10 g tryptone, 5 g yeast extract in 1 l of distilled water) with 2 ml autoclaved $MgSO_4$, 20 ml autoclaved solution 50xM (1.25 M Na_2HPO_4, 1.25 M KH_2PO_4, 2.5 M NH_4Cl, 0.25 M Na_2SO_3) and 20 ml filter-sterilised solution 50x5052 (0.5% (w/v) glycerol, 0.05% w/v) glucose, 0.2% (w/v) α-lactose monohydrate). ZYM-5052 medium was supplement with the appropriate selection antibiotics.

A 1-litre ZYM-5052 media flask was inoculated with 40 ml of overnight culture of BL21(DE3) cells, harbouring the expression vector, grown in LB media with 0.2% (w/v) glucose and the appropriate antibiotic.

Cells were first grown at 37°C for the first four hours and then culture was incubated overnight at room temperature. Presence of the protein was analysed by SDS-PAGE gel by comparing protein band pattern of the inoculum culture with the auto-induction culture. Cells after induction were harvested by centrifuging at 9000 x *g* for 30 min at 15°C. Aliquots of 10 g were stored at -80°C.

7.6 Protein purification

7.6.1 Ni-affinity purification of histidine-tagged proteins from the supernatant

Supernatant containing secreted histidine-tagged protein was filtered through a Whatmann paper to remove any particulates which could block the column. To reduce nonspecific binding to the column, 200 mM NaCl and 20 mM imidazole was added to supernatant. The pH was adjusted to pH 8 with 0.5 M NaOAc, pH 4, if required.

Ni-chelate agarose beads (Amocol) were loaded into the column and washed with 10 column volumes (CV) of distilled water. Column was washed with 5 CV of wash buffer (50 mM NaOAc, pH 3.7; 300 mM NaCl) to remove loosely bound nickel ions. The column was then equilibrated with 10 CV of buffer A (25 mM Tris-HCl, pH 8; 500 mM NaCl, 20 mM imidazole, pH 8; 5% (v/v) glycerol). Equilibrated Ni-chelate agarose beads were then transferred into a container with culture supernatant and were incubated for at least an hour with mixing. After incubation the beads were allowed to sediment and supernatant was carefully decanted. Remaining Ni-chelate resin was loaded into the column. Supernatant was passed through the column and the column was washed with 10 CV of buffer A. The histidine-tagged protein was eluted from the column with a 10 CV linear gradient of imidazole from 20-500 mM in the same buffer; 1 CV fractions were collected during the elution. Following a wash with high imidazole buffer (250 mM), the column was washed with 6 M guanidine hydrochloride with 10 mM EDTA to remove any tightly bound proteins and the nickel ions. Agarose beads were then washed with water and the column was recharged with 100 mM $NiCl_2$ in 20% (v/v) ethanol.

The eluted 1CV fractions were analysed by SDS-PAGE and fractions containing the histidine-tagged protein were pooled together, concentrated on VIVA spin column (Sartorius) with the membrane cut-off of 5 kDa and dialysed into 25 mM NaOAc, 50 mM NaCl, pH 4. Purified protein was stored at 4°C or flash frozen in liquid nitrogen and stored at -80°C.

7.6.2 Purification of proteins from the cell pellet

7.6.2.1 Crude cell extract preparations

Approximately 5 g of the induced cell paste was thawed on ice and then resuspended in 50 ml of lysis buffer (50 mM Tris-HCl, pH 8.2; 2 mM EDTA, 200 mM NaCl, 5% glycerol (v/v)). A freshly prepared solution of lysozyme was added to a final concentration of 200 µg/ml and the cell suspension was stirred on ice for approximately 30 min. A 2.3 mg/ml stock solution of phenylmethanesulfonyl fluoride (PMSF) in 100% ethanol was added to a final

concentration of 23 µg/ml followed by addition of sodium deoxycholate to a final concentration of 500 µg/ml and DTT to a final concentration of 5 mM. The cell lysate was stirred for 30 min on ice until it became viscous. The suspension was sonicated at 10 kHz and 28 µm amplitude three times for 30 seconds (5 sec pulse, 5 sec pause) and centrifuged at 40000 x g for 30 minutes at 4°C. Soluble and insoluble fractions were analysed by SDS-PAGE for the presence of native GlyΔSP.

7.6.2.2 Purification of histidine-tagged proteins from the soluble fraction on Ni-affinity column

Crude cell extract were prepared as described in Section 7.6.2.1; however, EDTA and DTT were not added to the cell lysate as they would interfere with Ni-affinity column. Cell lysates were spun down at 40000 x g for 30 min. The soluble faction was filtered through a Whatman paper filter to remove any precipitant and supernatant was diluted 1:1 with Buffer A (see Section 7.6.1). The protocol for protein purification on Ni-affinity column was followed as described in Section 7.6.1.

7.6.3 Ni-affinity purification of the histidine tagged proteins in denaturing conditions

N-histidine tagged EPB4.2 was purified on the Ni-affinity column under denaturing conditions due to its instability. Cell pellet was thawed on ice and resuspended in lysis buffer (50 mM Tris-HCl, pH8; 200 mM NaCl, 5% glucose, 5 mM imidazole) containing EDTA-free protease inhibitor (Thermo Scientific). Solid guanidine hydrochloride (GuHCl) was added to the suspension to the final concentration of approximately 6.5 M. The cell lysate was briefly sonicated to reduce viscosity and centrifuged at 40000 x g for 30 minutes at 4°C. The method for standard Ni-affinity chromatography described in Section 7.6.1 was followed with modifications. Clear cell lysate was passed through Ni-affinity column (Amocol) equilibrated in Buffer A (see Section 7.6.1) containing 3 M GuHCl. The column was washed with 5 CV of Buffer A + 3 M GuHCl. The

denaturing buffer was removed by the column wash with 10 CV of the Buffer A. EPB4.2 was eluted in linear gradient of imidazole ranging from 20 mM to 500 mM in the same buffer Fractions equal to one CV were continuously collected. Elution fractions were analysed by SDS-PAGE.

7.6.3.1 Gly1ORF1ΔSP purification from the insoluble fraction

Inclusion bodies isolation and resolubilisation

The crude cell lysate of the induced BL21(DE3):pET21-Gly1ΔSP was prepared as described in Section 7.6.2.1. The insoluble cell fraction was washed three times with 1% (w/v) sodium deoxycholate with 1 mM EDTA (10 ml per 1 g of initial cell pellet weight). Insoluble clumps were resuspended by brief sonication. The sodium deoxycholate was then removed by two washes with sterile distilled water. The purified inclusion bodies were stored at -20°C.

Approximately 20 mg of wet mass of inclusion bodies was solubilised in 1 ml of 8 M guanidine hydrochloride in 50 mM Tris-HCl, pH 8, 5 mM DTT for 2 hours at room temperature. For small scale refolding screen the DTT was removed using GE Healthcare PD Trap^{TM} G-25 desalting columns equilibrated with 8 M GuHCl in 50 mM Tris-HCl, pH 8. Samples before refolding trials were centrifuged for 20 min at 40000 x g to remove any aggregates that could form aggregation starting points during refolding.

Refolding screen

The Thermo Scientific Pierce Refolding Kit was used to determine the optimal conditions for protein refolding. As Gly1ORF1 has one disulphide bond it is expected that for correct refolding protein needs an optimal reducing environment. The manufacturer's protocol for the primary screen examining the effect on refolding of denaturant concentration, reducing environment (GSH:GSSG) and protein concentration was followed. Stocks of 1.1 x base refolding buffers contained 55 mM Tris, 21 mM NaCl, 0.88 mM KCl and varying concentrations of guanidine hydrochloride and L- arginine were prepared. To 900 µl those buffers were added 1 mM EDTA and different concentration of GSH and GSSG; the volume of the buffer was made up to 950 µl with distilled

water. Two stock concentrations of the protein were tested; 10 mg/ml and 1 mg/ml.

Denatured protein was added in 10 µl aliquots to the refolding buffers and mixed by immediate vortexing. At least 2 minutes were allowed between adding additional aliquots. Once the protein was added to all the buffers, refolding was allowed to proceed overnight at room temperature. The next day, samples were examined for presence of aggregation by measuring absorbance at 450 nm. Samples with little or no aggregation were dialysed in 50 mM Tris-HCl, 200 mM NaCl and 100 mM L-arginine, pH 8.2.

Separation of native Gly1ΔSP on Size exclusion Chromatography (SEC)

After overnight incubation in the refolding buffer the native Gly1ΔSP was concentrated using VivaSpin columns with the membrane cut-off of 5 kDa (Sartorius Staedim). A GE Healthcare Superose 6 10/300 GL column was equilibrated with separation buffer (50 mM Tris, 200 mM NaCl, 100 mM L-arginine, 1 mM EDTA, pH 8.2). The protein sample was loaded onto the column and eluted at 0.5 ml/min collecting 0.5 ml fractions. Fractions that showed detectable absorbance at UV 260 nm, were analysed by SDS-PAGE in reducing and non-reducing conditions.

7.6.4 Purification of proteins from the periplasmic fraction

A 1 g aliquot of induced cells was first washed with 5 ml of EDTA wash buffer (20 mM Tris-HCl, pH 8; 150 mM NaCl, 5 mM EDTA) for 15 min on the roller. The cells were centrifuged at 5000 x g for 10 min at 15 °C. The resulting pellet was resuspended in 4 ml of ice-cold sucrose solution (20 mM Tris-HCl, pH 8, 25% (w/v) sucrose, 5 mM EDTA) and incubated on ice for 15 min. Cells were centrifuged at 8500 x g for 10 min at 4°C and exposed to osmotic shock by adding 5 ml of ice-cold 4 mM $MgCl_2$ containing 0.023 µg/ml PMSF and then gently resuspended. Lysozyme was added to a final concentration of 150 µg/ml and the sample incubated for 30 min on ice.

7.6.4.1 Polyethylenamine (PEI) precipitation

First, 0.5 M of solid ammonium sulphate was added to the EDTA wash fraction and dissolved. Polyethylenimine (PEI) solution at the concentration of 5% (w/v) was then added to a final concentration of approximately 0.075% (w/v). The precipitation of nucleic acid was carried out for 30 minutes at 4°C and sample was centrifuged for 30 minutes at 40000 x g. Protein in the supernatant containing the neisserial Gly1ORF1 paralog was precipitated by adding solid ammonium sulphate to a final concentration of 3.5 M followed by stirring for 1 hour at 4°C. Precipitated proteins were pelleted by centrifugation at 40000 x g for 30 minutes. Pellet was resuspended in the buffer of the choice. Resuspended pellet and the supernatant were dialysed against the buffer of choice.

7.6.4.2 Ion exchange chromatography

Ion exchange chromatography was used to remove remaining contaminants after Ni-affinity purification of the histidine-tagged proteins and after purification of neisserial Gly1ORF1 paralog from the periplasmic fraction. The ion exchange chromatography was carried out using either anion exchange column (Hi-Trap Q) or cation exchange columns: (Hi-Trap SP and/or(Hi-Trap SP). Pre-packed columns (1 ml and 5 ml) were purchased from GE Healthcare. SP and heparin columns were equilibrated with KP buffer at the pH optimised experimentally for each protein (25 mM potassium phosphate, 100 mM NaCl, 5% (v/v) glycerol, 1 mM EDTA and 1 mM DTT). Q columns were equilibrated with Q buffer (25 mM Tris-HCl pH 8, 100 mM NaCl, 5% (v/v) glycerol, 1mM EDTA and 1mM DTT). Clear sample was loaded onto the equilibrated column. Proteins were eluted from the column with 20 CV of the appropriate buffer at the linear salt gradient ranging from 100 mM to 1 M NaCl. Fractions at a size of 1 CV were collected continuously and analysed by SDS-PAGE gel.

7.7 Electrophoresis methods

7.7.1 Separation of DNA using agarose gel electrophoresis

Plasmid DNA and PCR amplification products were separated using 0.8% - 1.5 % (w/v) agarose gels. Powdered agarose was dissolved by boiling in 1x TAE buffer. The solution was allowed to cool down to the temperature of approximately 55°C and ethidium bromide was added to a final concentration of 0.5 µg/ml. The agarose solution was poured in the appropriate plate and allowed to set with an inserted comb. The DNA separation was carried in the electrophoresis tank in the 1 x TAE buffer at a voltage of 1-5 V/cm. DNA bands were visualised using UV-trans-illuminator.

7.7.2 SDS-Polyacrylamide gel electrophoresis

A 10 ml resolving gel was prepared by mixing: 4.5 ml of distilled water, 3.3 ml 30% acrylamide (0.8% bisacrylamide), 2 ml 200 mM Tris/bicine pH 8.3/1 mM EDTA, 100 µl 10% (w/v) SDS, 54 µl ammonium persulphate (100 mg/ml) and 16 µl TEMED (N,N, N',N'-tetramethylethylenediamine). Stacking gels were prepared by mixing 0.5 ml 30% acrylamide (0.8% bisacrylamide), 1 ml 250 mM Tris-HCl pH 6.8, 1.5 ml distilled water, 30 µl of 10% (w/v) SDS, 15 µl ammonium persulphate (100 mg/ml) and 7.5 µl of TEMED. The stacking gel was poured on the top of solidified resolving gel and the comb forming wells was immediately inserted.

The protein samples were mixed 1:1 with 2xSDS loading buffer (90 mM Tris-HCl, pH 6.8; 20 mM EDTA, 35 % (v/v) glycerol, 0.05% (w/v) bromophenol blue, 1.8% (w/v) SDS, 1.3 mM DTT) and boiled for 3-5 min. The samples were loaded into the wells of the SDS-PAGE gel plates. Electrophoresis was carried out in the electrophoresis buffer (100 mM Tris-Bicine pH 8.3, 1 mM EDTA, 1% (w/v) SDS) initially at 20 mV and the voltage was increased to 40 mV after the proteins reached the resolving gel.

After completion of SDS-PAGE electrophoresis gels were stained in Coomassie blue stain comprising of 0.05% (w/v) Coomassie brilliant blue R-250 dissolved in the mixture of distilled water, methanol and glacial acetic acid in a ratio of

5:4:1. The stained gels were then destained in the destain solution comprising of distilled water, methanol and glacial acetic acid in a ratio of 7:2:1.

7.7.3 Separation of proteins on the native PAGE gel

Polyacrylamide gel was prepared as previously described in Section 7.7.2 but without the addition of SDS. Similarly the composition of the PAGE running buffer and 2 x sample loading buffer were the same as previously used except for the absence of SDS. Protein samples were mixed 1:1 with 2x sample loading buffer and the samples loaded onto the gel. The PAGE was run at 100 V for four hours at 4°C. The gels were stained with Coomasie stain and destained with destaining solution as described in Section 7.7.2.

7.7.4 Protein detection by western blot

The protein samples were first separated by SDS-PAGE as described in Section 7.7.2. The gel and a nitrocellulose membrane were soaked for 15 minutes in the transfer buffer (2.9 g Tris-HCl, 1 g of glycine and 100 ml of methanol, the volume was adjusted to 500 ml with sterile distilled water). The protein bands were transferred onto a cellulose membrane in the semi-dry transfer cell (BIO-RAD, CA, USA) at 10 V for 1 hour. After the transfer, the membrane was incubated in the blocking buffer (PBS with 0.05% (v/v) Tween-20 – PBS-T containing 5% (w/v) milk powder) overnight at 4°C or for one hour at room temperature. The membrane was washed with PBS-T at room temperature for 30 minutes and the membrane was probed with the primary antibody in PBS-T containing 1% (w/v) milk powder for one hour at room temperature. The excess of the antibody was removed by the wash step with PBS-T for 30 minutes. The membrane was than probed with appropriate secondary antibody conjugated with HRP for one hour at room temperature, followed by the wash step with PBST. The proteins were detected after adding the HRP substrate (Clarity Western ECL Substrate, BIO-RAD) and visualised on the X-ray film (KODAK).

7.8 Analytical methods for studying physico-biochemical properties of proteins

7.8.1 Haemin agarose pull-down assay

Haemin agarose pull down is a commonly used analytical tool to identify haem binding proteins. An aliquot of haemin agarose beads suspension (25 µl) (Sigma) was washed three times with 1 ml of PBS buffer to remove the storage buffer. A 0.5 ml aliquot of protein samples in PBS at the concentration of approximately 100 µg/mL containing BSA was mixed with the beads and incubated on the roller at room temperature for 30 min. Samples were centrifuged at 5000 x *g* for 2 min to sediment the beads and supernatant was removed. Beads were incubated in PBS with addition of 0.5 M NaCl for 15 min and then sedimented by centrifugation. Two more washes with PBS were performed and the sedimented beads were resuspended in 30 µl 2x SDS loading buffer and boiled for 5 minutes. Samples were then analysed by SDS-PAGE.

7.8.2 Haemin absorbance spectra in presence of Gly1ORF1

A 0.5 mM stock of monomeric haemin was prepared as described previously (Cuív *et al.*, 2008). Solid haemin (3.3 mg) was dissolved in 500 µl 0.1 M NaOH and the pH of the solution was reduced by addition of an equal volume of 1 M Tris-HCl, pH 8. Solubilised haem was diluted with 20 mM Tris-HCl, pH 7.5, 150 mM NaCl to a final volume of 10 ml. The concentration of the haemin stock was confirmed by measuring absorbance of the stock at 385 nm and calculating the actual concentration using an extinction coefficient (ε) of 58400 $M^{-1}cm^{-1}$. Gly1ORF solution was prepared in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl to a final concentration of 5 µM and this was mixed with an increasing concentration of haemin. Spectra were recorded from 260 nm to 600 mM with 5 nm steps.

7.8.3 Microscale thermophoresis (MST)

Protein labelling

Proteins for analysis were fluorescently labelled using the NanoTemper Monolith RED labelling kit according to the manufacturer's instruction. Briefly, proteins were diluted to a concentration of 10 µM and then buffer exchanged into PBS using Zeba desalting spin columns (Thermo Scientific). The dye was dissolved in DMSO and diluted in PBS to the final concentration of 20 µM. The dye was mixed with an equal volume of the protein solution resulting in a dye:protein ratio of 2:1. The reaction was then incubated in the dark at room temperature for 1 hour. Uncoupled dye was removed from the solution using PD MiniTrap^{TM} G-25 columns and proteins eluted in 500 µl of MST buffer provided by NanoTemper Technologies GmbH (50 mM Tris-HCl pH 7.6, 150 mM NaCl, 10 mM $MgCl_2$, 0.05% (v/v) Tween-20). The samples were then stored in aliquots at -80°C until further analysis.

Gly1ORF1 – haemin binding analysis by microscale thermophoresis (MST)

Protein samples were first centrifuged for 5 min at 15000 x *g*. Haemin was dissolved in DMSO and diluted in MST to a final concentration of 100 diluted in MST to a final concentration of 100 µM. Protein concentration was kept constant at 0.2 µM throughout the experiment. Haemin concentrations ranged from 50 µM to 0.181 µM in serial 3:2 dilution steps in MST buffer + 1% (v/v) DMSO to avoid the dilution effect of DMSO during experiment. Serial dilutions of haemin were made in NanoTemper tubes to a final volume of 10 µl and an equal volume of protein solution at 400 nM were added to each tube. Samples were left for 5 minutes at room temperature to equilibrate and then loaded into MST NT.115 hydrophilic glass capillaries. The measurements were taken using a Monolith NT.115 instrument (NanoTemper Technologies GmbH) at 25°C with the LED and MST powers set at 20%. Data were analysed using the NT Analysis Software (NanoTemper Technologis GmbH).

Two models were used for data fitting:

1. K_d formula (law of mass action):

$$F_{NORM}(C_T) = F_{NORM,F} + \frac{F_{NORM,B} - F_{NORM,F}}{2}\left(C_F + C_T + K_d - \sqrt{(C_F + C_T + K_d)^2 - 4C_F C_T}\right)$$

2. Hill equation

$$F_{NORM}(C_T) = F_{NORM,F} + \frac{F_{NORM,B} - F_{NORM,F}}{1 + \left(\frac{EC_{50}}{C_T}\right)^n}$$

F_{NORM} - normalised fluorescence, measured for the respective titrant concentration (C_T)

$F_{NORM,F}$ - normalised fluorescence of the unbound (free) state of the fluorescently labelled molecules,

$F_{NORM,B}$ - normalised fluorescence of the bound state of the fluorescently labelled molecules,

C_F- concentration of the fluorescently labelled molecules (constant),

C_T- concentration of the titrant,

K_d - equilibrium dissociation constant,

EC_{50} - half maximal effective concentration (concentration at which half of the fluorescently

labelled molecules are bound),

n - Hill coefficient.

7.8.4 Dynamic Light Scattering

Dynamic light scattering was carried out to determine the hydrodynamic size of Gly1ORF1 in the solution and to determine the distribution of the population of its multimeric forms. For the experiments, Gly1ORF1 C-his was dialysed against 25 mM NaOAC, 50 mM NaCl, pH 4.

The sterile filtered solution at 0.5 mg/ml concentration was analysed in triplicate on Malvern Zetasizer Nano Spectrophotometer, Serial number MAL1057018 (The Department of Chemistry, University of Sheffield).

To check if multimeric forms could be disrupted in the reducing condition, DTT was added at different concentration and sample was incubated on ice for 10 minutes before taking the measurement.

7.8.5 Analysis of oligomerisation by cross-linking approach

Gly1ORF1 was previously shown to exhibit multimeric forms due to intermolecular disulphide bond formation. This multimerisation can be visualised on SDS PAGE gel under non-reducing conditions. To force the formation of multimers, samples were treated with a non-reversible cross-linker - glutaraldehyde. All Gly1ORF1 mutants created in this study were tested for their levels of multimerisation to check if any of the mutations causes structural changes which could disrupt the intermolecular interaction between Gly1 monomers.

Gly1ORF1 was diluted in PBS to a final concentration of 100 µg/ml. Aliquots of Gly1ORF1 mutant solution were mixed in a ratio 1:1 with varying concentration of glutaraldehyde diluted in PBS. The oligomerisation reaction was carried for one hour and reaction stopped by adding 1 M Tris-HCl, pH 8 to a final concentration of 50 mM and incubation was carried out for another 15 minutes. Samples were analysed by western blot using anti-Gly1ORF1 antibodies and appropriate HRP-labelled antibody.

7.9 Methods for studying Gly1ORF1 interactions with red blood cells

7.9.1 Preparation of red blood cell suspension.

Blood from the healthy volunteer was collected by venepuncture into heparinised sterile container after written consent had been obtained (University of Sheffield Ethical Approval No SMBRER39).

An aliquot of 5 ml of fresh blood was washed three times with ten volumes of PBS and the red blood cells (RBC) were resuspended in PBS to a haematocrit of approximately 10%.

7.9.2 Preparation of red blood cell ghosts

Ghosts were obtained by lysing RBC in hypotonic conditions and removing the haemoglobin. Resuspending membranes in PBS resulted in formation of empty red blood cell membrane vesicles.

Freshly drawn blood in heparin was washed three times with ten volumes of PBS. Red blood cells were resuspended in 10 ml of sterile distilled water and incubated for 15 minutes at room temperature with gentle rolling. Lysed RBCs were spun down for 10 minutes at 13000 x *g* and membranes were washed three times or until white with distilled water for 15 minutes. Ghosts were resuspended in PBS and stored at 4°C.

7.9.3 Cross-linking of red blood cells with Gly1ORF1

A 50 μl aliquots of red blood cell suspension from a healthy volunteer resuspended in PBS to a 10% haematocrit were mixed with an equal volume of Gly1 at 100 μg/ml in PBS with addition the of 400 μg/ml of BSA as a carrier. Samples were incubated for one hour either at 37°C or 4°C. Control with Gly1ORF1 without red blood cells was also prepared in parallel. After pre-probing of Gly1 with red blood cells, 100 μl of glutaraldehyde dilutions in PBS were added resulting in samples with final glutaraldehyde concentration of 0, 0.05, 0.25, 0.5, 1 and 2%. The cross linking reaction was allowed to proceed for one hour under the same conditions as the pre-probing with Gly1. The reaction was stopped by adding 1 M Tris-HCl, pH 8 to a final concentration of 20 mM Tris-HCl.

The interaction of Gly1 with RBCs and ghost were also analysed using reversible cross-linkers: 3,3'-dithiobis[sulfosuccinimidylpropionate] – DTSSP, anddimethyl 3,3'-dithiobispropionimidate*2HCl – DTBP. DTBP is membrane permeable, whereas DTSSP can be only used for cross-linking on the cell surface. Stock concentrations of DTBP and DTSSP were prepared in PBS buffer and added to the reaction to a required final concentration. The reactions were carried out as described above. The cross-linked samples were analysed in the non-reducing SDS-PAGE gel.

7.9.4 Analysis of dissociation of Gly1ORF1 from red blood cells

A 50 µl aliquot of red blood cell suspension was incubated with 50 µl of Gly1ORF1 C-histidine tagged at 100 µg/ml in PBS with BSA at 400 µg/ml for one hour at room temperature with gentle rolling. After incubation red blood cells were gently spun down at 100 x g for 30 sec and the supernatant was removed. Red blood cells were then washed with multiple washes with 200 µl of PBS for 15 minutes. After each wash and resuspending the RBCs in fresh buffer, a 10 µl sample of the suspension was analysed by western blot using anti-Gly1ORF polyclonal antibodies.

7.9.5 FACS analysis of protein interactions with red blood cells

Red blood cells prepared as described in Section 7.9.1 were washed four times in ten volumes of PBS and diluted in PBS to a haematocrit of approximately 0.5 %. A 250 µl aliquot of diluted RBCs in PBS was mixed with 250 µl of 100 µg/ml Gly1ORF1 in PBS containing four times molar excess of BSA. This was incubated overnight at 4°C. Negative control with RBCs mixed with PBS was carried out in parallel. After overnight incubation RBCs were washed twice with PBS.

The samples were probed with primary rabbit anti-Gly1ORF1 antibodies affinity purified in PBS containing 1.5% (w/v) of BSA at a 1:100 dilution for 1 hour at room temperature followed by two washes with PBS and 20 min probing at 4°C with FITC – labelled anti-rabbit secondary antibody (BD Biosciences) at a 1:250 dilution. Samples were washed twice in PBS and resuspended in the same buffer to the final volume of 500 µl. Samples were analysed on FACS Calibur at the Flow-Cytometry Core Facility, University of Sheffield

Alternatively, a FITC-labelled protein was incubated with red blood cells and incubated in the conditions described above. The red blood cells were then washed four times with PBS and analysed on FACS Calibur.

7.10 Methods for studying haemoglobin receptor HmbR interactions

7.10.1 Congo red binding plates

The method was based on previously described protocol (Liu *et al.*, 2012a) with modifications. Cells grown in LB media containing 0.1% (w/v) glucose and ampicillin were grown to mid log phase and then streaked out on BHI plate containing 0.005% (w/v) Congo red, 0.05% (w/v) α-mono lactose and ampicillin at 100 µg.ml and grown overnight. at 30°C. BL21 Star™ (DE3) cells with an empty pET21 vector were used as a negative control.

7.10.2 Congo red and haemin binding analysis by SDS-PAGE gel staining

Cell lysates of induced and control cells were separated by SDS PAGE gel described previously (see Section 7.7.2). After SDS-PAGE, gels were washed first with TBG-buffer (100 mM TRIS, 100 mM Bicine, 50 mM NaCl, 10 mM MgCl$_2$, KCl, 10% (v/v) glycerol; pH 8.3) containing 5 mM DTT for 20 minutes followed by two washes with TBG buffer without DTT. Ability to bind Congo red and haemin was determined by incubating the gels with TBG buffer containing 80 µg/ml of Congo red or 100 µM haemin in TBG buffer with no DTT. The gels were destained by several washes with TBG buffer.

7.10.3 Binding of haem and haemoglobin by the *E. coli* cells expressing HmbR

In order to establish if expressed HmbR is present on the surface of the cells, a whole cell binding assay was performed with *E. coli* BL21 Star™ (DE3) overproducing the protein of interests as described previously (Simpson *et al.*, 2000).

Freshly transformed *E. coli* BL21 Star™ (DE3) cells with an empty pET21 vector and a vector containing *hmbR* gene were inoculated into LB media containing 0.1% (w/v) glucose and ampicillin and incubated overnight at 37°C. The overnight cultures were used as inoculates for 100 ml cultures containing glucose and ampicillin. Cultures were grown at 30°C until they reached an

$OD_{600\ nm}$ of approximately 0.6 and were induced by adding 0.1 mM IPTG. Aliquots of induced cultures were taken at 0, 30, 60, 90 and 120 minutes post induction. Cultures were pelleted at 50000 x g and resuspended in PBS to an $OD_{600\ nm}$ of 1. Aliquots of 0.8 ml of cell suspensions were transferred into sterile tubes and 0.2 ml of human haemoglobin and haemin solutions in PBS were added to a final concentration of 10 µM and 5 µM, respectively. Cells were incubated with haem and haemoglobin for an hour at 37°C and centrifuged at 5000 x g. The absorbance of the supernatants was measured on the spectrophotometer at 385 nm (haemin) and at 410 nm (haemoglobin).

7.10.4 Binding of Gly1ORF1 to *E. coli* cells expressing HmbR

Cell cultures of induced HmbR and control with an empty pET21 vector were prepared as described in Section 7.10.3. Cell suspensions of un-induced and induced cultures were incubated with Gly1ORF1 at a final concentration of 2.5 µM at 37°V for 1 hour. The cells were centrifuged at 5000 x g for 3 minutes. Binding of Gly1ORF1 to the cells was determined by western blot comparing amount of Gly1ORF1 in supernatants with control sample with no added cells. Gly1ORF1 bound to the *E. coli* cells was also detected by western blot in the cell lysates. Images were analysed by ImageJ software (Schneider *et al.*, 2012).

7.11 Methods for evaluation of immunogenic properties of the proteins

7.11.1 Analysing antibody titres of mouse anti-Gly1ORF1 serum by the enzyme-linked immuno-absorbent assay (ELISA)

Mouse sera against Gly1ORF1 protein were raised and provided by BioServ UK Ltd., Sheffield. Six mice were injected with soluble Gly1ORF1 protein followed by a booster vaccination after four weeks. Control mice were injected with PBS. Sera from individual mice showed cross-reactivity with Gly1ORF1 when tested by western blot by Dr R. Walton.

ELISA method

Sera from six mice were pooled and antibody titre was determined by ELISA. 96-well ELISA plates were coated with 50 µl of Gly1ORF1 in PBS at the concentration of 5 µg/ml and then incubated overnight at 4°C. Plates coated with antigen were washed three times with PBST (PBS containing 0.05% (v/v) Tween 20). Individual wells were blocked with 100 µl of blocking buffer (PBST containing 5% (w/v) skimmed milk powder) for 1 hour at room temperature with shaking. The blocking step was followed by three washes with PBST. Twelve, two-fold dilutions of mouse sera in blocking buffer starting from 1:80 were added to sequential wells in duplicate. Plates were incubated for 1 hour at room temperature with shaking and then washed three times with PBST buffer. Anti-mouse IgG – HRP secondary antibody (Cell Signalling) in blocking buffer at 1:10000 dilution were then added and incubation was carried out for another hour at room temperature with shaking. After washing it three times with PBST, 80 µl of substrate for HRP ABTS (2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) was added. The reaction was stopped after 10 minutes by adding equal volume of 1% (w/v) SDS in each well and plates were read on the plate reader at 405 nm.

Controls in this assay included wells without antigen incubated with first dilution of sera to test unspecific interactions and antigen incubated only with secondary antibody.

7.11.2 Detection of antibody response against Gly1ORF1 in human sera

Sera form 6 patients with acute meningococcal infection taken during admission to the hospital and one month after infection were obtained as a kind gift from Prof. RC. Read. The 96 sera samples from patients with no suspected meningococcal infection from the Serum Epidemiological Unit (SEU) were obtained from the Health Protection Agency. The method described in Section 7.11.1 was followed with slight modifications.

To reduce the background, the Tween 20 in blocking buffer and in the wash buffer was increased to 0.1% (w/v) concentration and skimmed milk powder

was replaced with BSA 3% (w/v). The secondary antibody used in this assay was goat HRP-conjugated anti-human IgG (alpha chain) antibody (Sigma). Sera from 96 patients from SEU were initially tested at a 1:20 dilution with three repeats. Samples with the highest anti-Gly1ORF1 levels were re-tested at two-fold serial dilutions.

7.11.3 Measuring of Serum Bactericidal Activities of anti-Gly1ORF1 antibodies

7.11.3.1 Preparation of the human complement for SBA

Freshly drawn blood from a healthy volunteer obtained under written consent (local ethics permission, SMBRER39) was transferred into a sterile tube and left on the bench to clot for 30 min. The tube containing clotted blood was then placed on ice and incubated for an hour. A blood clot was pelleted by centrifugation at 800 x g for 10 min at 4°C and serum was carefully aspirated to a fresh tube. Serum was then incubated for an hour on ice with suspension of *N. meningitidis* MC58 prepared in PBS. This procedure removed the antibodies cross-reacting with NM MC58 and reduced the background killing of the complement. Bacteria were removed by centrifugation and human complement was filter-sterilised, dispensed into aliquots and stored at -80°C until needed.

7.11.3.2 Affinity purification of anti-Gly1ORF1 rabbit polyclonal antibodies.

CNBr-activated Sepharose 4B beads (1 g, Sigma) were washed and swollen with 200 ml of ice cold 1 mM HCl on a sintered glass filter (porosity G3). The beads were rapidly washed three times with the coupling buffer (0.1 M carbonate buffer, 0.5 M NaCl, pH 8.3) in a 50 ml polypropylene tube and centrifuged for 2-3 min at 50 x g after each wash. Approximately 10 mg of protein in 10 ml of coupling buffer was added to 1 ml of washed beads. Tubes were incubated overnight at 4°C on the rotating mixer. Samples were taken immediately at the start and at end of incubation to determine the coupling efficiency.

Beads were centrifuged for 2 min at 70 x *g* and washed twice with coupling buffer, then resuspended with 15 ml of 0.1 M Tris-HCl, pH 8.0 to inactivate any remaining active groups on the beads. Beads were incubated for 2 hours at room temperature on a shaker. Three washes at 4°C were performed, first with 0.1 M Tris-HCl, pH 8 followed with 0.1 M acetate buffer, pH 4.0. Finally, beads were washed three times with PBS.

Gly1ORF1-coupled beads (1 ml) were resuspended in 10 ml PBS and an equal volume of rabbit serum raised against Gly1ORF1 was added. Binding of antiserum to the beads was performed overnight at 4°C on a rotating mixer. Elution of the antibodies was performed in the cold room. First, the beads were washed three times with PBS and then resuspended in 10 ml of PBS. The beads were loaded into a disposable plastic column and washed approximately with 10 column volumes of PBS or until no protein was detected spectrophotometrically at 280 nm.

Bound antibodies were eluted with 200 mM glycine, pH 2.8 and 1 ml fractions were collected into tubes containing 27 µl of 3 M Tris-HCl, pH 8.8 and 100 µl of 3 M KCl. Purified antibodies were dialysed against PBS with three buffer changes. The dialysed antibodies were concentrated, filter-sterilised and aliquots were stored at -80°C.

7.11.3.3   SBA – serum bactericidal assay

N. meningitidis MC58 WT and isogenic Gly1ORF1 knock-out bacteria were grown for 16-18 h on Columbia Blood Agar (CBA) plates at 37°C in an atmosphere of 5% CO_2. Isogenic Gly1ORF1 mutant NM MC58 was grown on plates supplemented with 5 µg/ml of erythromycin. A loopful of bacteria was transferred from the plate into 2 ml of PBSB (PBS with 0.9 mM $CaCl_2$ 0.5 mM $MgCl_2$) containing 2% (v/v) de-complemented foetal calf serum (FCS). Cells were gently resuspended and centrifuged for 2 min at 200 x *g* to remove any bacterial aggregates. Supernatant containing homogenous suspension of bacteria was transferred to a clean tube.

Polyclonal sera and FCS were de-complemented by incubation at 56°C for 30 min. Bacterial suspension containing approximately 1000 cfu (25 µl) was added

to the wells of a sterile 96-well culture plate. Serial dilutions of de-complemented serum were prepared in PBS and added to the bacterial suspensions. Human serum as an exogenous complement source was freshly thawed and added to a final concentration of 5% (v/v). The volume was made up to 100 μl with PBSB containing 2% (v/v) FCS. Plates were agitated briefly and incubated for 30 min at 37°C with 5% (v/v) CO_2. After incubation, 15 μl from each well was plated on CBA plates in triplicate to determine viable count. Controls included samples with bacteria only, bacteria with human complement and bacteria with heat-inactivated human complement. The first dilution of polyclonal antibodies with heat inactivated complement was also tested. The percentage of killing was calculated by using control sample with bacteria and human complement as a 100% and calculating percentage difference in cfu in tested samples.

7.11.4 Whole Cell ELISA with *Neisseria* spp.

Coating of the ELISA microtitre plates with suspension of homogenised bacteria was performed as previously described (Abdillahi and Poolman, 1987). Different strains of *Neisseria* were grown on CBA plates for 16-18 hours at 37°C with 5% CO_2. Bacteria were scraped from the plates with a sterile swab and resuspended in PBS. The bacterial suspension was adjusted to an optical density of OD_{600} 0.1. Bacterial suspension was pasteurised for 30 min at 56°C in a water bath and aliquots of 100 μl were transferred into individual wells of ELISA microtitre plates. Plates were incubated uncovered overnight at 37°C to allow evaporation of the buffer and fixation of the cells.

Plates were then washed three times with deionised water containing 0.025% (v/v) Tween 20. Primary antibodies were diluted in blocking buffer (PBS, 3% (w/v) BSA, 0.01% Tween 20) and probed with immobilised cells for one hour at 37°C on the shaker. Plates were then washed three times with deionised water with 0.025% (v/v) Tween and secondary antibody was added at a 1:1000 dilution in blocking buffer. Plates were incubated for 1 h at room temperature with gentle shaking. After washing the plates three times with deionised water with 0.025% (v/v) Tween 20, a substrate for HRP (ABTS) was added and the

colour development was allowed to proceed for 20 min and absorbance at 405 nm was measured on the plate reader.

7.11.5 Detection of the antigen on the surface of the *Neisseria* using flow cytometry

A protocol for the procedure was adapted from Hung and co-workers (Hung *et al.*, 2013). Analysed strains were grown overnight at 37°C on CBA plates in an atmosphere of 5% (v/v) CO_2. Cells were scraped from the plates and resuspended in PBS. Cells were washed two times with PBS and resuspended in 2 ml of ice cold 70% (v/v) ethanol and incubated for one hour at -20°C. This step allowed permeabilisation of the bacterial capsule and allowed interaction of the surface antigens with the primary antibody. Cells were then washed three times with PBS containing 1% (w/v) BSA and resuspended in PBS + 1% (BSA) to a density of approximately 10^6 cfu/ml of suspension. Bacterial suspensions were aliqouted into sterile tubes (final volume of 0.5 ml) and affinity purified rabbit antibodies against Gly1ORF1 and alpha protein were added to the final dilution of 1:100. Cells were incubated with primary antibodies for 30 min at 37°C. Negative controls were incubated with PBS + 1% (w/v) BSA only.

7.12 Gly1ORF1 homologs bioinformatic analysis

The protein sequence of the neisserial Gly1ORF1 was submitted to blastp with PSI-blast algorithm (NCBI). The hits from the blast search were further iterated. From a number of PSI-blast results following protein sequences has been chosen: hypothetical protein Hflu203001319 (Ref. ID ZP_00157632.2) of *Haemophilus influenzae* R2866 and hypothetical protein mh0603 (Ref. ID ABG89179.1) of *Mannheimia haemolytica*

The physico-chemical properties of the protein were analysed by ProtParam Expasy Proteomics Server (URL: http://expasy.org/tools/protparam.html) and the search for motifs and conserved domains was done with the InterProScan server (URL: http://www.ebi.ac.uk/Tools/InterProScan/). The subcellular

localisation of the hypothetical proteins was determined with the PSORTb version 3.0.2 (URL: http://www.psort.org/psortb/). The site of the cleavage of the N-terminal signal peptide was predicted by SignalP 3.0 software (URL: http://www.cbs.dtu.dk/services/SignalP/). To predict the secondary structure topology the GLY1ORF1 homolog sequences were submitted to SABLE server (URL: http://sable.cchmc.org/), whereas the disulphide bonds formation probability was analysed with the help of DISULFIND server (URL: http://disulfind.dsi.unifi.it/).

REFERENCES

Abdillahi, H., and Poolman, J.T. (1987). Whole-cell ELISA for typing Neisseria meningitidis with monoclonal antibodies. FEMS Microbiology Letters *48*, 367-371.

Achaz, G., Rocha, E.P.*, et al.* (2002). Origin and fate of repeats in bacteria. Nucleic Acids Res *30*, 2987-2994.

Adamczak, R., Porollo, A.*, et al.* (2005). Combining prediction of secondary structure and solvent accessibility in proteins. Proteins *59*, 467-475.

Ala'aldeen, D.A., Flint, M.*, et al.* (2010). Human antibody responses to the meningococcal factor H binding protein (LP2086) during invasive disease, colonization and carriage. Vaccine *28*, 7667-7675.

Alamro, M., Bidmos, F.A.*, et al.* (2014). Phase variation mediates reductions in expression of surface proteins during persistent meningococcal carriage. Infect Immun *82*, 2472-2484.

Altschul, S.F., Madden, T.L.*, et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402.

Andrews, S.M., and Pollard, A.J. (2014). A vaccine against serogroup B Neisseria meningitidis: dealing with uncertainty. Lancet Infect Dis *14*, 426-434.

Anjum, M.F., Stevanin, T.M.*, et al.* (2002). Nitric oxide metabolism in Neisseria meningitidis. J Bacteriol *184*, 2987-2993.

Armbruster, C.E., Smith, S.N.*, et al.* (2014). Increased incidence of urolithiasis and bacteremia during Proteus mirabilis and Providencia stuartii coinfection due to synergistic induction of urease activity. J Infect Dis *209*, 1524-1532.

Arnold, R., Galloway, Y.*, et al.* (2011). Effectiveness of a vaccination programme for an epidemic of meningococcal B in New Zealand. Vaccine *29*, 7100-7106.

Arvidson, C.G., Kirkpatrick, R., *et al.* (1999). Neisseria gonorrhoeae mutants altered in toxicity to human fallopian tubes and molecular characterization of the genetic locus involved. Infect Immun *67*, 643-652.

Arvidson, D.N., Pearson, R.F., *et al.* (2003). Purification, characterization and preliminary X-ray crystallographic studies on Neisseria gonorrhoeae Gly1ORF1. Acta Crystallogr D Biol Crystallogr *59*, 747-748.

Ayala, B.P., Vasquez, B., *et al.* (2001). The pilus-induced Ca2+ flux triggers lysosome exocytosis and increases the amount of Lamp1 accessible to Neisseria IgA1 protease. Cell Microbiol *3*, 265-275.

Bachman, M.A., Miller, V.L., *et al.* (2009). Mucosal lipocalin 2 has pro-inflammatory and iron-sequestering effects in response to bacterial enterobactin. PLoS Pathog *5*, e1000622.

Baker, H.M., and Baker, E.N. (2012). A structural perspective on lactoferrin function. Biochem Cell Biol *90*, 320-328.

Beck, S.C., and Meyer, T.F. (2000). IgA1 protease from Neisseria gonorrhoeae inhibits TNFalpha-mediated apoptosis of human monocytic cells. FEBS Lett *472*, 287-292.

Bernard, S.C., Simpson, N., *et al.* (2014). Pathogenic Neisseria meningitidis utilizes CD147 for vascular colonization. Nat Med *20*, 725-731.

Bernhard, S., Spaniol, V., *et al.* (2012). Molecular pathogenesis of infections caused by Moraxella catarrhalis in children. Swiss Med Wkly *142*, w13694.

Bernheimer, A.W. (1988). Assay of hemolytic toxins. Methods Enzymol *165*, 213-217.

Bhattacharyya, R., Das, A.K., *et al.* (1999). Mapping of a palmitoylatable band 3-binding domain of human erythrocyte membrane protein 4.2. Biochem J *340 (Pt 2)*, 505-512.

Bjelland, S., and Seeberg, E. (2003). Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. Mutat Res *531*, 37-80.

Bjune, G., Høiby, E.A., *et al.* (1991). Effect of outer membrane vesicle vaccine against group B meningococcal disease in Norway. Lancet *338*, 1093-1096.

Black, C.G., Fyfe, J.A.*, et al.* (1998). Absence of an SOS-like system in Neisseria gonorrhoeae. Gene *208*, 61-66.

Blom, A.M., Hallström, T.*, et al.* (2009). Complement evasion strategies of pathogens-acquisition of inhibitors and beyond. Mol Immunol *46*, 2808-2817.

Borrow, R., Aaberge, I.S.*, et al.* (2005a). Interlaboratory standardization of the measurement of serum bactericidal activity by using human complement against meningococcal serogroup b, strain 44/76-SL, before and after vaccination with the Norwegian MenBvac outer membrane vesicle vaccine. Clin Diagn Lab Immunol *12*, 970-976.

Borrow, R., Balmer, P.*, et al.* (2005b). Meningococcal surrogates of protection--serum bactericidal antibody activity. Vaccine *23*, 2222-2227.

Boukhalfa, H., and Crumbliss, A.L. (2002). Chemical aspects of siderophore mediated iron transport. Biometals *15*, 325-339.

Bove, P.F., and van der Vliet, A. (2006). Nitric oxide and reactive nitrogen species in airway epithelial signaling and inflammation. Free Radic Biol Med *41*, 515-527.

Bracken, C.S., Baer, M.T.*, et al.* (1999). Use of heme-protein complexes by the Yersinia enterocolitica HemR receptor: histidine residues are essential for receptor function. J Bacteriol *181*, 6063-6072.

Braconier, J.H., Sjöholm, A.G.*, et al.* (1983). Fulminant meningococcal infections in a family with inherited deficiency of properdin. Scand J Infect Dis *15*, 339-345.

Brekke, O.L., Hellerud, B.C.*, et al.* (2011). Neisseria meningitidis and Escherichia coli are protected from leukocyte phagocytosis by binding to erythrocyte complement receptor 1 in human blood. Mol Immunol *48*, 2159-2169.

Brown, D.R., Helaine, S.*, et al.* (2010). Systematic functional analysis reveals that a set of seven genes is involved in fine-tuning of the multiple functions mediated by type IV pili in Neisseria meningitidis. Infect Immun *78*, 3053-3063.

Bruce, L.J. (2008). Red cell membrane transport abnormalities. Curr Opin Hematol *15*, 184-190.

Butera, D., Cook, K.M.*, et al.* (2014). Control of blood proteins by functional disulfide bonds. Blood *123*, 2000-2007.

Callaghan, M.J., Jolley, K.A.*, et al.* (2006). Opacity-associated adhesin repertoire in hyperinvasive Neisseria meningitidis. Infect Immun *74*, 5085-5094.

Calmettes, C., Alcantara, J.*, et al.* (2012). The structural basis of transferrin sequestration by transferrin-binding protein B. Nat Struct Mol Biol *19*, 358-360.

Carpenter, E.P., Corbett, A.*, et al.* (2007). AP endonuclease paralogues with distinct activities in DNA repair and bacterial pathogenesis. EMBO J *26*, 1363-1372.

Carson, S.D., Klebba, P.E.*, et al.* (1999). Ferric enterobactin binding and utilization by Neisseria gonorrhoeae. J Bacteriol *181*, 2895-2901.

Carson, S.D., Stone, B.*, et al.* (2000). Phase variation of the gonococcal siderophore receptor FetA. Mol Microbiol *36*, 585-593.

Carter, D.M., Miousse, I.R.*, et al.* (2006). Interactions between TonB from Escherichia coli and the periplasmic protein FhuD. J Biol Chem *281*, 35413-35424.

Cartwright, K.A., Jones, D.M.*, et al.* (1991). Influenza A and meningococcal disease. Lancet *338*, 554-557.

Cassat, J.E., and Skaar, E.P. (2013). Iron in infection and immunity. Cell Host Microbe *13*, 509-519.

Ceroni, A., Passerini, A.*, et al.* (2006). DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. Nucleic Acids Res *34*, W177-181.

Chacon-Cruz, E., Espinosa-De Los Monteros, L.E.*, et al.* (2014). An outbreak of serogroup C (ST-11) meningococcal disease in Tijuana, Mexico. Ther Adv Vaccines *2*, 71-76.

Chen, C.J., Tobiason, D.M.*, et al.* (2004). A mutant form of the Neisseria gonorrhoeae pilus secretin protein PilQ allows increased entry of heme and antimicrobial compounds. J Bacteriol *186*, 730-739.

Clarke, T.E., Tari, L.W.*, et al.* (2001). Structural biology of bacterial iron uptake systems. Curr Top Med Chem *1*, 7-30.

Cornelissen, C.N., and Hollander, A. (2011). TonB-Dependent Transporters Expressed by Neisseria gonorrhoeae. Front Microbiol *2*, 117.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res *16*, 10881-10890.

Coureuil, M., Bourdoulous, S.*, et al.* (2014). Invasive meningococcal disease: a disease of the endothelial cells. Trends Mol Med.

Coureuil, M., Join-Lambert, O.*, et al.* (2013). Pathogenesis of meningococcemia. Cold Spring Harb Perspect Med *3*.

Criss, A.K., and Seifert, H.S. (2012). A bacterial siren song: intimate interactions between Neisseria and neutrophils. Nat Rev Microbiol *10*, 178-190.

Cuív, P.O., Keogh, D.*, et al.* (2008). The hmuUV genes of Sinorhizobium meliloti 2011 encode the permease and ATPase components of an ABC transport system for the utilization of both haem and the hydroxamate siderophores, ferrichrome and ferrioxamine B. Mol Microbiol *70*, 1261-1273.

Czjzek, M., Létoffé, S.*, et al.* (2007). The crystal structure of the secreted dimeric form of the hemophore HasA reveals a domain swapping with an exchanged heme ligand. J Mol Biol *365*, 1176-1186.

Dahl, K.N., Parthasarathy, R.*, et al.* (2004). Protein 4.2 is critical to CD47-membrane skeleton attachment in human red cells. Blood *103*, 1131-1136.

Darton, T.C., Jack, D.L.*, et al.* (2014). MBL2 Deficiency is Associated with Higher Genomic Bacterial Loads during Meningococcemia in Young Children. Clin Microbiol Infect.

Das, A.K., Bhattacharya, R.*, et al.* (1994). Human erythrocyte membrane protein 4.2 is palmitoylated. Eur J Biochem *224*, 575-580.

Daskaleros, P.A., and Payne, S.M. (1987). Congo red binding phenotype is associated with hemin binding and increased infectivity of Shigella flexneri in the HeLa cell model. Infect Immun *55*, 1393-1398.

Davidsen, T., Bjørås, M._, et al._ (2005). Antimutator role of DNA glycosylase MutY in pathogenic Neisseria species. J Bacteriol _187_, 2801-2809.

Davidsen, T., and Tønjum, T. (2006). Meningococcal genome dynamics. Nat Rev Microbiol _4_, 11-22.

Davidsen, T., Tuven, H.K._, et al._ (2007). Genetic interactions of DNA repair pathways in the pathogen Neisseria meningitidis. J Bacteriol _189_, 5728-5737.

De Domenico, I., Zhang, T.Y._, et al._ (2010). Hepcidin mediates transcriptional changes that modulate acute cytokine-induced inflammatory responses in mice. J Clin Invest _120_, 2395-2405.

Deghmane, A.E., Larribe, M._, et al._ (2003). Differential expression of genes that harbor a common regulatory element in Neisseria meningitidis upon contact with target cells. Infect Immun _71_, 2897-2901.

Dehio, C., Gray-Owen, S.D._, et al._ (1998). The role of neisserial Opa proteins in interactions with host cells. Trends Microbiol _6_, 489-495.

Delany, I., Grifantini, R._, et al._ (2006). Effect of Neisseria meningitidis fur mutations on global control of gene transcription. J Bacteriol _188_, 2483-2492.

Deneer, H.G., and Potter, A.A. (1989). Iron-repressible outer-membrane proteins of Pasteurella haemolytica. J Gen Microbiol _135_, 435-443.

DeVoe, I.W. (1976). Egestion of degraded meningococci by polymorphonuclear leukocytes. J Bacteriol _125_, 258-266.

Diaz Romero, J., and Outschoorn, I.M. (1994). Current status of meningococcal group B vaccine candidates: capsular or noncapsular? Clin Microbiol Rev _7_, 559-575.

Dietrich, G., Kurz, S._, et al._ (2003). Transcriptome analysis of Neisseria meningitidis during infection. J Bacteriol _185_, 155-164.

Doughari, H.J., Ndakidemi, P.A._, et al._ (2011). The ecology, biology and pathogenesis of Acinetobacter spp.: an overview. Microbes Environ _26_, 101-112.

Doulet, N., Donnadieu, E.*, et al.* (2006). Neisseria meningitidis infection of human endothelial cells interferes with leukocyte transmigration by preventing the formation of endothelial docking structures. J Cell Biol *173*, 627-637.

Drogari-Apiranthitou, M., Kuijper, E.J.*, et al.* (2002). Complement activation and formation of the membrane attack complex on serogroup B Neisseria meningitidis in the presence or absence of serum bactericidal activity. Infect Immun *70*, 3752-3758.

Dunn, K.L., Farrant, J.L.*, et al.* (2003). Bacterial [Cu,Zn]-cofactored superoxide dismutase protects opsonized, encapsulated Neisseria meningitidis from phagocytosis by human monocytes/macrophages. Infect Immun *71*, 1604-1607.

Eason, M.M., and Fan, X. (2014). The role and regulation of catalase in respiratory tract opportunistic bacterial pathogens. Microb Pathog *74C*, 50-58.

Echenique-Rivera, H., Muzzi, A.*, et al.* (2011). Transcriptome analysis of Neisseria meningitidis in human whole blood and mutagenesis studies identify virulence factors involved in blood survival. PLoS Pathog *7*, e1002027.

Elkins, C., Carbonetti, N.H.*, et al.* (1992). Antibodies to N-terminal peptides of gonococcal porin are bactericidal when gonococcal lipopolysaccharide is not sialylated. Mol Microbiol *6*, 2617-2628.

Escolar, L., de Lorenzo, V.*, et al.* (1997). Metalloregulation in vitro of the aerobactin promoter of Escherichia coli by the Fur (ferric uptake regulation) protein. Mol Microbiol *26*, 799-808.

Evans, C.M., Pratt, C.B.*, et al.* (2011). Nasopharyngeal colonization by Neisseria lactamica and induction of protective immunity against Neisseria meningitidis. Clin Infect Dis *52*, 70-77.

Evans, N.J., Harrison, O.B.*, et al.* (2010). Variation and molecular evolution of HmbR, the Neisseria meningitidis haemoglobin receptor. Microbiology *156*, 1384-1393.

Fang, F.C. (2004). Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. Nat Rev Microbiol *2*, 820-832.

Ferrari, G., Garaguso, I.*, et al.* (2006). Outer membrane vesicles from group B Neisseria meningitidis delta gna33 mutant: proteomic and immunological comparison with detergent-derived outer membrane vesicles. Proteomics *6*, 1856-1866.

Figueroa, J.E., and Densen, P. (1991). Infectious diseases associated with complement deficiencies. Clin Microbiol Rev *4*, 359-395.

Fillat, M.F. (2014). The FUR (ferric uptake regulator) superfamily: diversity and versatility of key transcriptional regulators. Arch Biochem Biophys *546*, 41-52.

Finne, J., Finne, U*., et al.* (1983). Occurrence of alpha 2-8 linked polysialosyl units in a neural cell adhesion molecule. Biochem Biophys Res Commun *112*, 482-487.

Frasch, C.E., Borrow, R*., et al.* (2009). Bactericidal antibody is the immunologic surrogate of protection against meningococcal disease. Vaccine *27 Suppl 2*, B112-116.

Frasch, C.E., Preziosi, M.P*., et al.* (2012). Development of a group A meningococcal conjugate vaccine, MenAfriVac(TM). Hum Vaccin Immunother *8*, 715-724.

Friedberg EC, W.G., Siede W, Wood RD, Schultz RA, Ellenberger T. (2006). DNA Repair and Mutagenesis. (Washington, D. C.: ASM Press).

Friedrichs, B., Koob, R*., et al.* (1989). Demonstration of immunoreactive forms of erythrocyte protein 4.2 in nonerythroid cells and tissues. Eur J Cell Biol *48*, 121-127.

Fusco, W.G., Choudhary, N.R*., et al.* (2013). Mutational analysis of hemoglobin binding and heme utilization by a bacterial hemoglobin receptor. J Bacteriol *195*, 3115-3123.

Geoffroy, M.C., Floquet, S*., et al.* (2003). Large-scale analysis of the meningococcus genome by gene disruption: resistance to complement-mediated lysis. Genome Res *13*, 391-398.

Geörg, M., Maudsdotter, L*., et al.* (2013). Meningococcal resistance to antimicrobial peptides is mediated by bacterial adhesion and host cell RhoA and Cdc42 signalling. Cell Microbiol *15*, 1938-1954.

Giuntini, S., Vu, D.M*., et al.* (2013). fH-dependent complement evasion by disease-causing meningococcal strains with absent fHbp genes or frameshift mutations. Vaccine *31*, 4192-4199.

Goldschneider, I., Gotschlich, E.C.*, et al.* (1969). Human immunity to the meningococcus. I. The role of humoral antibodies. J Exp Med *129*, 1307-1326.

Granick, S., and Beale, S.I. (1978). Hemes, chlorophylls, and related compounds: biosynthesis and metabolic regulation. Adv Enzymol Relat Areas Mol Biol *46*, 33-203.

Grifantini, R., Sebastian, S.*, et al.* (2003). Identification of iron-activated and - repressed Fur-dependent genes by transcriptome analysis of Neisseria meningitidis group B. Proc Natl Acad Sci U S A *100*, 9542-9547.

Grosdidier, A., Zoete, V.*, et al.* (2011a). Fast docking using the CHARMM force field with EADock DSS. J Comput Chem.

Grosdidier, A., Zoete, V.*, et al.* (2011b). SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Res *39*, W270-277.

Hagen, T.A., and Cornelissen, C.N. (2006). Neisseria gonorrhoeae requires expression of TonB and the putative transporter TdfF to replicate within cervical epithelial cells. Mol Microbiol *62*, 1144-1157.

Hamilton, H.L., and Dillard, J.P. (2006). Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. Mol Microbiol *59*, 376-385.

Haralambous, E., Dolly, S.O.*, et al.* (2006). Factor H, a regulator of complement activity, is a major determinant of meningococcal disease susceptibility in UK Caucasian patients. Scand J Infect Dis *38*, 764-771.

Harrison, O.B., Bennett, J.S.*, et al.* (2013). Distribution and diversity of the haemoglobin-haptoglobin iron-acquisition systems in pathogenic and non-pathogenic Neisseria. Microbiology *159*, 1920-1930.

Harrison, O.B., Evans, N.J.*, et al.* (2009). Epidemiological evidence for the role of the hemoglobin receptor, hmbR, in meningococcal virulence. J Infect Dis *200*, 94-98.

Hauck, C.R., and Meyer, T.F. (1997). The lysosomal/phagosomal membrane protein h-lamp-1 is a target of the IgA1 protease of Neisseria gonorrhoeae. FEBS Lett *405*, 86-90.

Hermeling, S., Crommelin, D.J.*, et al.* (2004). Structure-immunogenicity relationships of therapeutic proteins. Pharm Res *21*, 897-903.

Hill, D.J., Griffiths, N.J.*, et al.* (2010). Cellular and molecular biology of Neisseria meningitidis colonization and invasive disease. Clin Sci (Lond) *118*, 547-564.

Hollander, A., Mercante, A.D.*, et al.* (2011). The iron-repressed, AraC-like regulator MpeR activates expression of fetA in Neisseria gonorrhoeae. Infect Immun *79*, 4764-4776.

Hrkal, Z., Vodrázka, Z.*, et al.* (1974). Transfer of heme from ferrihemoglobin and ferrihemoglobin isolated chains to hemopexin. Eur J Biochem *43*, 73-78.

Huang, B. (2009). MetaPocket: a meta approach to improve protein ligand binding site prediction. OMICS *13*, 325-330.

Hung, M.C., Heckels, J.E.*, et al.* (2013). The adhesin complex protein (ACP) of Neisseria meningitidis is a new adhesin with vaccine potential. mBio *4*.

Imhaus, A.F., and Duménil, G. (2014). The number of Neisseria meningitidis type IV pili determines host cell interaction. EMBO J *33*, 1767-1783.

Izadi, N., Henry, Y.*, et al.* (1997). Purification and characterization of an extracellular heme-binding protein, HasA, involved in heme iron acquisition. Biochemistry *36,* 7050-7057.

Jabado, N., Jankowski, A.*, et al.* (2000). Natural resistance to intracellular infections: natural resistance-associated macrophage protein 1 (Nramp1) functions as a pH-dependent manganese transporter at the phagosomal membrane. J Exp Med *192*, 1237-1248.

Jack, D.L., Dodds, A.W.*, et al.* (1998). Activation of complement by mannose-binding lectin on isogenic mutants of Neisseria meningitidis serogroup B. J Immunol *160*, 1346-1353.

Jack, D.L., Jarvis, G.A.*, et al.* (2001). Mannose-binding lectin accelerates complement activation and increases serum killing of Neisseria meningitidis serogroup C. J Infect Dis *184*, 836-845.

Jafri, R.Z., Ali, A.*, et al.* (2013). Global epidemiology of invasive meningococcal disease. Popul Health Metr *11*, 17.

James, K.J., Hancock, M.A.*, et al.* (2009). TonB interacts with BtuF, the Escherichia coli periplasmic binding protein for cyanocobalamin. Biochemistry *48*, 9212-9220.

Jamet, A., Rousseau, C.*, et al.* (2009). A two-component system is required for colonization of host cells by meningococcus. Microbiology *155*, 2288-2295.

Jarva, H., Ram, S.*, et al.* (2005). Binding of the complement inhibitor C4bp to serogroup B Neisseria meningitidis. J Immunol *174*, 6299-6307.

Jennings, M.P., Srikhanta, Y.N.*, et al.* (1999). The genetic basis of the phase variation repertoire of lipopolysaccharide immunotypes in Neisseria meningitidis. Microbiology *145 (Pt 11),* 3013-3021.

Jones, A., Geörg, M.*, et al.* (2009). Endotoxin, capsule, and bacterial attachment contribute to Neisseria meningitidis resistance to the human antimicrobial peptide LL-37. J Bacteriol *191*, 3861-3868.

Jordan, P.W., and Saunders, N.J. (2009). Host iron binding proteins acting as niche indicators for Neisseria meningitidis. PLoS One *4*, e5198.

Jordens, J.Z., Williams, J.N.*, et al.* (2004). Development of immunity to serogroup B meningococci during carriage of Neisseria meningitidis in a cohort of university students. Infect Immun *72*, 6503-6510.

Jose, J., Wölk, U.*, et al.* (2000). Human T-cell response to meningococcal immunoglobulin A1 protease associated alpha-proteins. Scand J Immunol *51*, 176-185.

Kanzaki, A., Yawata, Y.*, et al.* (1995). Band 4.2 Komatsu: 523 GAT-->TAT (175 Asp-->Tyr) in exon 4 of the band 4.2 gene associated with total deficiency of band 4.2, hemolytic anemia with ovalostomatocytosis and marked disruption of the cytoskeletal network. Int J Hematol *61*, 165-178.

Kelley, L.A., and Sternberg, M.J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc *4*, 363-371.

Khavrutskii, L., Yeh, J.*, et al.* (2013). Protein purification-free method of binding affinity determination by microscale thermophoresis. J Vis Exp.

Kim, D.D., and Song, W.C. (2006). Membrane complement regulatory proteins. Clin Immunol *118*, 127-136.

Korsgren, C., Lawler, J.*, et al.* (1990). Complete amino acid sequence and homologies of human erythrocyte membrane protein band 4.2. Proc Natl Acad Sci U S A *87*, 613-617.

Krewulak, K.D., and Vogel, H.J. (2008). Structural biology of bacterial iron uptake. Biochim Biophys Acta *1778*, 1781-1804.

Kristiansen, P.A., Diomandé, F.*, et al.* (2013). Impact of the serogroup A meningococcal conjugate vaccine, MenAfriVac, on carriage and herd immunity. Clin Infect Dis *56*, 354-363.

Kumar, P., Sannigrahi, S.*, et al.* (2012). The Neisseria meningitidis ZnuD zinc receptor contributes to interactions with epithelial cells and supports heme utilization when expressed in Escherichia coli. Infect Immun *80*, 657-667.

Lappann, M., Haagensen, J.A.*, et al.* (2006). Meningococcal biofilm formation: structure, development and phenotypes in a standardized continuous flow system. Mol Microbiol *62*, 1292-1309.

Lappann, M., Otto, A.*, et al.* (2013). Comparative proteome analysis of spontaneous outer membrane vesicles and purified outer membranes of Neisseria meningitidis. J Bacteriol *195*, 4425-4435.

Larkin, M.A., Blackshields, G.*, et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Laver, J.R., Stevanin, T.M.*, et al.* (2010). Bacterial nitric oxide detoxification prevents host cell S-nitrosothiol formation: a novel mechanism of bacterial pathogenesis. FASEB J *24*, 286-295.

Lee, S.W., Higashi, D.L.*, et al.* (2005). PilT is required for PI(3,4,5)P3-mediated crosstalk between Neisseria gonorrhoeae and epithelial cells. Cell Microbiol *7*, 1271-1284.

Létoffé, S., Deniau, C.*, et al.* (2001). Haemophore-mediated bacterial haem transport: evidence for a common or overlapping site for haem-free and haem-loaded haemophore on its specific outer membrane receptor. Mol Microbiol *41*, 439-450.

Létoffé, S., Redeker, V._, et al._ (1998). Isolation and characterization of an extracellular haem-binding protein from Pseudomonas aeruginosa that shares function and sequence similarities with the Serratia marcescens HasA haemophore. Mol Microbiol *28*, 1223-1234.

Lewis, L.A., Gipson, M._, et al._ (1999). Phase variation of HpuAB and HmbR, two distinct haemoglobin receptors of Neisseria meningitidis DNM2. Mol Microbiol *32*, 977-989.

Lewis, L.A., Ngampasutadol, J._, et al._ (2010). The meningococcal vaccine candidate neisserial surface protein A (NspA) binds to factor H and enhances meningococcal resistance to complement. PLoS Pathog *6*, e1001027.

Lewis, L.A., and Ram, S. (2014). Meningococcal disease and the complement system. Virulence *5*, 98-126.

Lewis, L.A., Vu, D.M._, et al._ (2013). Factor H-Dependent Alternative Pathway Inhibition Mediated by Porin B Contributes to Virulence of Neisseria meningitidis. mBio *4*.

Lin, T.Y., Chan, M.C._, et al._ (2014). Clinical manifestations and prognostic factors of Morganella morganii bacteremia. Eur J Clin Microbiol Infect Dis.

Liu, M., Ferrandez, Y._, et al._ (2012a). Heme binding proteins of Bartonella henselae are required when undergoing oxidative stress during cell and flea invasion. PLoS One *7*, e48408.

Liu, X., Follmer, D._, et al._ (2012b). Characterization of the function of cytoglobin as an oxygen-dependent regulator of nitric oxide concentration. Biochemistry *51*, 5072-5082.

Liu, X., Olczak, T._, et al._ (2006). Identification of amino acid residues involved in heme binding and hemoprotein utilization in the Porphyromonas gingivalis heme receptor HmuR. Infect Immun *74*, 1222-1232.

Lo, H., Tang, C.M._, et al._ (2009). Mechanisms of avoidance of host immunity by Neisseria meningitidis and its effect on vaccine development. Lancet Infect Dis *9*, 418-427.

Lorenzen, D.R., Düx, F._, et al._ (1999). Immunoglobulin A1 protease, an exoenzyme of pathogenic Neisseriae, is a potent inducer of proinflammatory cytokines. J Exp Med *190*, 1049-1058.

Lovett, S.T. (2004). Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences. Mol Microbiol *52*, 1243-1253.

Lowell, G.H., Smith, L.F.*, et al.* (1979). Antibody-dependent cell-mediated antibacterial activity of human mononuclear cells. I. K lymphocytes and monocytes are effective against meningococi in cooperation with human imune sera. J Exp Med *150*, 127-137.

Lucidarme, J., Tan, L.*, et al.* (2011). Characterization of Neisseria meningitidis isolates that do not express the virulence factor and vaccine antigen factor H binding protein. Clin Vaccine Immunol *18*, 1002-1014.

Macindoe, G., Mavridis, L.*, et al.* (2010). HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic Acids Res *38*, W445-449.

Mackie, J.T., Barton, M.*, et al.* (1995). Pasteurella haemolytica septicaemia in sheep. Aust Vet J *72*, 474.

Madico, G., Welsch, J.A.*, et al.* (2006). The meningococcal vaccine candidate GNA1870 binds the complement regulatory protein factor H and enhances serum resistance. J Immunol *177*, 501-510.

Malik, S., Sami, M.*, et al.* (1993). A role for band 4.2 in human erythrocyte band 3 mediated anion transport. Biochemistry *32*, 10078-10084.

Malorny, B., Morelli, G.*, et al.* (1998). Sequence diversity, predicted two-dimensional protein structure, and epitope mapping of neisserial Opa proteins. J Bacteriol *180*, 1323-1330.

Mandrell, R.E., Kim, J.J.*, et al.* (1991). Endogenous sialylation of the lipooligosaccharides of Neisseria meningitidis. J Bacteriol *173*, 2823-2832.

McNeil, L.K., Murphy, E.*, et al.* (2009). Detection of LP2086 on the cell surface of Neisseria meningitidis and its accessibility in the presence of serogroup B capsular polysaccharide. Vaccine *27*, 3417-3421.

Meadows, C. (2004). Characterisation of Emi protein and alpha peptides of *Neisseria*

meningitidis. . (University of Sheffield).

Merz, A.J., Enns, C.A.*, et al.* (1999). Type IV pili of pathogenic Neisseriae elicit cortical plaque formation in epithelial cells. Mol Microbiol *32*, 1316-1332.

Merz, A.J., Rifenbery, D.B.*, et al.* (1996). Traversal of a polarized epithelium by pathogenic Neisseriae: facilitation by type IV pili and maintenance of epithelial barrier function. Mol Med *2*, 745-754.

Merz, A.J., and So, M. (1997). Attachment of piliated, Opa- and Opc- gonococci and meningococci to epithelial cells elicits cortical actin rearrangements and clustering of tyrosine-phosphorylated proteins. Infect Immun *65*, 4341-4349.

Merz, A.J., So, M.*, et al.* (2000). Pilus retraction powers bacterial twitching motility. Nature *407*, 98-102.

Moran, A.P., Prendergast, M.M.*, et al.* (1996). Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease. FEMS Immunol Med Microbiol *16*, 105-115.

Morgan, B.P., Marchbank, K.J.*, et al.* (2005). Complement: central to innate immunity and bridging to adaptive responses. Immunol Lett *97*, 171-179.

Morgenthau, A., Beddek, A.*, et al.* (2014). The negatively charged regions of lactoferrin binding protein B, an adaptation against anti-microbial peptides. PLoS One *9*, e86243.

Morgenthau, A., Livingstone, M.*, et al.* (2012). The role of lactoferrin binding protein B in mediating protection against human lactoferricin. Biochem Cell Biol *90*, 417-423.

Nagorska, K., Silhan, J.*, et al.* (2012). A network of enzymes involved in repair of oxidative DNA damage in Neisseria meningitidis. Mol Microbiol *83*, 1064-1079.

Nairz, M., Schroll, A.*, et al.* (2010). The struggle for iron - a metal at the host-pathogen interface. Cell Microbiol *12*, 1691-1702.

Netto, L.E., and Stadtman, E.R. (1996). The iron-catalyzed oxidation of dithiothreitol is a biphasic process: hydrogen peroxide is involved in the initiation of a free radical chain of reactions. Arch Biochem Biophys *333*, 233-242.

Noinaj, N., Buchanan, S.K.*, et al.* (2012a). The transferrin-iron import system from pathogenic Neisseria species. Mol Microbiol *86*, 246-257.

Noinaj, N., Easley, N.C., *et al.* (2012b). Structural basis for iron piracy by pathogenic Neisseria. Nature *483*, 53-58.

Noinaj, N., Guillier, M., *et al.* (2010). TonB-dependent transporters: regulation, structure, and function. Annu Rev Microbiol *64*, 43-60.

Nørskov-Lauritsen, N., Bruun, B., *et al.* (2012). Identification of haemolytic Haemophilus species isolated from human clinical specimens and description of Haemophilus sputorum sp. nov. Int J Med Microbiol *302*, 78-83.

O'Connor, C.D., and Timmis, K.N. (1987). Highly repressible expression system for cloning genes that specify potentially toxic proteins. J Bacteriol *169*, 4457-4462.

Oster, P., Lennon, D., *et al.* (2005). MeNZB: a safe and highly immunogenic tailor-made vaccine against the New Zealand Neisseria meningitidis serogroup B disease epidemic strain. Vaccine *23*, 2191-2196.

Pacher, P., Beckman, J.S., *et al.* (2007). Nitric oxide and peroxynitrite in health and disease. Physiol Rev *87*, 315-424.

Parker Siburt, C.J., Mietzner, T.A., *et al.* (2012). FbpA--a bacterial transferrin with more to offer. Biochim Biophys Acta *1820*, 379-392.

Parsons, H. (2003). The role of the alpha-peptide subunit of IGS protease in the pathogenesis of meningococcal disease. . (University of Sheffield).

Perkins-Balding, D., Baer, M.T., *et al.* (2003). Identification of functionally important regions of a haemoglobin receptor from Neisseria meningitidis. Microbiology *149*, 3423-3435.

Perkins-Balding, D., Ratliff-Griffin, M., *et al.* (2004). Iron transport systems in Neisseria meningitidis. Microbiol Mol Biol Rev *68*, 154-171.

Pettersson, A., Maas, A., *et al.* (1995). Molecular characterization of FrpB, the 70-kilodalton iron-regulated outer membrane protein of Neisseria meningitidis. Infect Immun *63*, 4181-4184.

Peyssonnaux, C., Zinkernagel, A.S.*, et al.* (2006). TLR4-dependent hepcidin expression by myeloid cells in response to bacterial pathogens. Blood *107*, 3727-3732.

Posey, J.E., and Gherardini, F.C. (2000). Lack of a role for iron in the Lyme disease pathogen. Science *288*, 1651-1653.

Post, D.M., Zhang, D.*, et al.* (2005). Biochemical and functional characterization of membrane blebs purified from Neisseria meningitidis serogroup B. J Biol Chem *280*, 38383-38394.

Proft, T., and Baker, E.N. (2009). Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease. Cell Mol Life Sci *66*, 613-635.

Puklo, M., Guentsch, A.*, et al.* (2008). Analysis of neutrophil-derived antimicrobial peptides in gingival crevicular fluid suggests importance of cathelicidin LL-37 in the innate immune response against periodontogenic bacteria. Oral Microbiol Immunol *23*, 328-335.

Ram, S., Sharma, A.K.*, et al.* (1998). A novel sialic acid binding site on factor H mediates serum resistance of sialylated Neisseria gonorrhoeae. J Exp Med *187*, 743-752.

Rock, J.D., and Moir, J.W. (2005). Microaerobic denitrification in Neisseria meningitidis. Biochem Soc Trans *33*, 134-136.

Rohde, K.H., and Dyer, D.W. (2004). Analysis of haptoglobin and hemoglobin-haptoglobin interactions with the Neisseria meningitidis TonB-dependent receptor HpuAB by flow cytometry. Infect Immun *72*, 2494-2506.

Roos, D., van Bruggen, R.*, et al.* (2003). Oxidative killing of microbes by neutrophils. Microbes Infect *5*, 1307-1315.

Rosenstein, N.E., Perkins, B.A.*, et al.* (2001). Meningococcal disease. N Engl J Med *344*, 1378-1388.

Rossi, E. (2005). Hepcidin--the iron regulatory hormone. Clin Biochem Rev *26*, 47-49.

Rossi, M.S., Fetherston, J.D.*, et al.* (2001). Identification and characterization of the hemophore-dependent heme acquisition system of Yersinia pestis. Infect Immun *69*, 6707-6717.

Roussel-Jazédé, V., Arenas, J.*, et al.* (2014). Variable processing of the IgA protease autotransporter at the cell surface of Neisseria meningitidis. Microbiology.

Roussel-Jazédé, V., Grijpstra, J.*, et al.* (2013). Lipidation of the autotransporter NalP of Neisseria meningitidis is required for its function in the release of cell-surface-exposed proteins. Microbiology *159*, 286-295.

Rudel, T., Scheurerpflug, I.*, et al.* (1995). Neisseria PilC protein identified as type-4 pilus tip-located adhesin. Nature *373*, 357-359.

Runyen-Janecky, L.J. (2013). Role and regulation of heme iron acquisition in gram-negative pathogens. Front Cell Infect Microbiol *3*, 55.

Saliba, A.E., Vonkova, I.*, et al.* (2014). A quantitative liposome microarray to systematically characterize protein-lipid interactions. Nat Methods *11*, 47-50.

Satchwell, T.J., Shoemark, D.K.*, et al.* (2009). Protein 4.2: a complex linker. Blood Cells Mol Dis *42*, 201-210.

Sathyamurthy, S. (2011). The role of Gly1, a secreted neisserial protein, in meningococcal pathogenesis. (University of Sheffield).

Saunders, N.J., Jeffries, A.C.*, et al.* (2000). Repeat-associated phase variable genes in the complete genome sequence of Neisseria meningitidis strain MC58. Mol Microbiol *37*, 207-215.

Sayers, J.R., Krekel, C.*, et al.* (1992). Rapid high-efficiency site-directed mutagenesis by the phosphorothioate approach. Biotechniques *13*, 592-596.

Schneider, C.A., Rasband, W.S.*, et al.* (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675.

Schneidman-Duhovny, D., Inbar, Y.*, et al.* (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res *33*, W363-367.

Schneidman-Duhovny, D., Inbar, Y*., et al.* (2003). Taking geometry to its edge: fast unbound rigid (and hinge-bent) docking. Proteins *52*, 107-112.

Schoen, C., Joseph, B*., et al.* (2007). Living in a changing environment: insights into host adaptation in Neisseria meningitidis from comparative genomics. Int J Med Microbiol *297*, 601-613.

Scholten, R.J., Kuipers, B*., et al.* (1994). Lipo-oligosaccharide immunotyping of Neisseria meningitidis by a whole-cell ELISA with monoclonal antibodies. J Med Microbiol *41*, 236-243.

Seib, K.L., Serruto, D*., et al.* (2009). Factor H-binding protein is important for meningococcal survival in human whole blood and serum and in the presence of the antimicrobial peptide LL-37. Infect Immun *77*, 292-299.

Seiler, F., Lepper, P.M*., et al.* (2014). Regulation and function of antimicrobial peptides in immunity and diseases of the lung. Protein Pept Lett *21*, 341-351.

Serruto, D., Bottomley, M.J*., et al.* (2012). The new multicomponent vaccine against meningococcal serogroup B, 4CMenB: immunological, functional and structural characterization of the antigens. Vaccine *30 Suppl 2*, B87-97.

Siburt, C.J., Roulhac, P.L*., et al.* (2009). Hijacking transferrin bound iron: protein-receptor interactions involved in iron transport in N. gonorrhoeae. Metallomics *1*, 249-255.

Silva, L.P., Yu, R.H*., et al.* (2012). Steric and allosteric factors prevent simultaneous binding of transferrin-binding proteins A and B to transferrin. Biochem J *444*, 189-197.

Simpson, W., Olczak, T*., et al.* (2000). Characterization and expression of HmuR, a TonB-dependent hemoglobin receptor of Porphyromonas gingivalis. J Bacteriol *182*, 5737-5748.

Smalley, J.W., Birss, A.J*., et al.* (1995). Kinetics of Congo-red binding by haemin-limited and haemin-excess cells of Porphyromonas gingivalis W50. Anaerobe *1*, 201-207.

Smith, A.D., and Wilks, A. (2012). Extracellular heme uptake and the challenges of bacterial cell membranes. Curr Top Membr *69*, 359-392.

Smith, I., Caugant, D.A.*, et al.* (2006). High case-fatality rates of meningococcal disease in Western Norway caused by serogroup C strains belonging to both sequence type (ST)-32 and ST-11 complexes, 1985-2002. Epidemiol Infect *134*, 1195-1202.

Snyder, L.A., Davies, J.K.*, et al.* (2005). Comparative overview of the genomic and genetic differences between the pathogenic Neisseria strains and species. Plasmid *54*, 191-218.

Sotolongo, F., Campa, C.*, et al.* (2007). Cuban Meningococcal BC Vaccine: Experiences & Contributions from 20 Years of Application. MEDICC Rev *9*, 16-22.

Stanwell-Smith, R.E., Stuart, J.M.*, et al.* (1994). Smoking, the environment and meningococcal disease: a case control study. Epidemiol Infect *112*, 315-328.

Stephens, D.S. (2009). Biology and pathogenesis of the evolutionarily successful, obligate human bacterium Neisseria meningitidis. Vaccine *27 Suppl 2*, B71-77.

Stevanin, T.M., Laver, J.R.*, et al.* (2007). Metabolism of nitric oxide by Neisseria meningitidis modifies release of NO-regulated cytokines and chemokines by human macrophages. Microbes Infect *9*, 981-987.

Stevanin, T.M., Moir, J.W.*, et al.* (2005). Nitric oxide detoxification systems enhance survival of Neisseria meningitidis in human macrophages and in nasopharyngeal mucosa. Infect Immun *73*, 3322-3329.

Stojiljkovic, I., Larson, J.*, et al.* (1996). HmbR outer membrane receptors of pathogenic Neisseria spp.: iron-regulated, hemoglobin-binding proteins with a high level of primary structure conservation. J Bacteriol *178*, 4670-4678.

Stork, M., Bos, M.P.*, et al.* (2010). An outer membrane receptor of Neisseria meningitidis involved in zinc acquisition with vaccine potential. PLoS Pathog *6*, e1000969.

Strange, H.R., Zola, T.A.*, et al.* (2011). The fbpABC operon is required for Ton-independent utilization of xenosiderophores by Neisseria gonorrhoeae strain FA19. Infect Immun *79*, 267-278.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr Purif *41*, 207-234.

Stugard, C.E., Daskaleros, P.A.*, et al.* (1989). A 101-kilodalton heme-binding protein associated with congo red binding and virulence of Shigella flexneri and enteroinvasive Escherichia coli strains. Infect Immun *57*, 3534-3539.

Su, Y., Ding, Y.*, et al.* (2006). Associations of protein 4.2 with band 3 and ankyrin. Mol Cell Biochem *289*, 159-166.

Sung, L.A., Chien, S.*, et al.* (1990). Molecular cloning of human protein 4.2: a major component of the erythrocyte membrane. Proc Natl Acad Sci U S A *87*, 955-959.

Talà, A., Monaco, C.*, et al.* (2011). Glutamate utilization promotes meningococcal survival in vivo through avoidance of the neutrophil oxidative burst. Mol Microbiol *81*, 1330-1342.

Tan, L.K., Carlone, G.M.*, et al.* (2010). Advances in the development of vaccines against Neisseria meningitidis. N Engl J Med *362*, 1511-1520.

Tauseef, I., Ali, Y.M.*, et al.* (2013). Phase variation of PorA, a major outer membrane protein, mediates escape of bactericidal antibodies by Neisseria meningitidis. Infect Immun *81*, 1374-1380.

Tauseef, I., Harrison, O.B.*, et al.* (2011). Influence of the combination and phase variation status of the haemoglobin receptors HmbR and HpuAB on meningococcal virulence. Microbiology *157*, 1446-1456.

Thompson, S.A., and Sparling, P.F. (1993). The RTX cytotoxin-related FrpA protein of Neisseria meningitidis is secreted extracellularly by meningococci and by HlyBD+ Escherichia coli. Infect Immun *61*, 2906-2911.

Thompson, S.A., Wang, L.L.*, et al.* (1993). Neisseria meningitidis produces iron-regulated proteins related to the RTX family of exoproteins. J Bacteriol *175*, 811-818.

Tolosano, E., and Altruda, F. (2002). Hemopexin: structure, function, and regulation. DNA Cell Biol *21*, 297-306.

Toye, A.M., Ghosh, S.*, et al.* (2005). Protein-4.2 association with band 3 (AE1, SLCA4) in Xenopus oocytes: effects of three natural protein-4.2 mutations associated with hemolytic anemia. Blood *105*, 4088-4095.

Trotter, C., Findlow, J.*, et al.* (2007a). Seroprevalence of bactericidal and anti-outer membrane vesicle antibodies to Neisseria meningitidis group B in England. Clin Vaccine Immunol *14*, 863-868.

Trotter, C.L., Chandra, M.*, et al.* (2007b). A surveillance network for meningococcal disease in Europe. FEMS Microbiol Rev *31*, 27-36.

Tsumoto, K., Umetsu, M.*, et al.* (2004). Role of arginine in protein refolding, solubilization, and purification. Biotechnol Prog *20*, 1301-1308.

Turner, P.C., Thomas, C.E.*, et al.* (2001). Neisserial TonB-dependent outer-membrane proteins: detection, regulation and distribution of three putative candidates identified from the genome sequences. Microbiology *147*, 1277-1290.

Tzeng, Y.L., Ambrose, K.D.*, et al.* (2005). Cationic antimicrobial peptide resistance in Neisseria meningitidis. J Bacteriol *187*, 5387-5396.

Tzeng, Y.L., Zhou, X.*, et al.* (2006). Autoregulation of the MisR/MisS two-component signal transduction system in Neisseria meningitidis. J Bacteriol *188*, 5055-5065.

van Putten, J.P., and Paul, S.M. (1995). Binding of syndecan-like cell surface proteoglycan receptors is required for Neisseria gonorrhoeae entry into human mucosal cells. EMBO J *14*, 2144-2154.

Vidal, S., Tremblay, M.L.*, et al.* (1995). The Ity/Lsh/Bcg locus: natural resistance to infection with intracellular parasites is abrogated by disruption of the Nramp1 gene. J Exp Med *182*, 655-666.

Virji, M. (1997). Post-translational modifications of meningococcal pili. Identification of common substituents: glycans and alpha-glycerophosphate--a review. Gene *192*, 141-147.

Virji, M. (2009). Pathogenic neisseriae: surface modulation, pathogenesis and infection control. Nat Rev Microbiol *7*, 274-286.

Virji, M., Alexandrescu, C.*, et al.* (1992a). Variations in the expression of pili: the effect on adherence of Neisseria meningitidis to human epithelial and endothelial cells. Mol Microbiol *6*, 1271-1279.

Virji, M., Makepeace, K.*, et al.* (1993). Meningococcal Opa and Opc proteins: their role in colonization and invasion of human epithelial and endothelial cells. Mol Microbiol *10*, 499-510.

Virji, M., Makepeace, K.*, et al.* (1992b). Expression of the Opc protein correlates with invasion of epithelial and endothelial cells by Neisseria meningitidis. Mol Microbiol *6*, 2785-2795.

Virji, M., Makepeace, K.*, et al.* (1995). Opc- and pilus-dependent interactions of meningococci with human endothelial cells: molecular mechanisms and modulation by surface polysaccharides. Mol Microbiol *18*, 741-754.

Virji, M., Watt, S.M.*, et al.* (1996). The N-domain of the human CD66a adhesion molecule is a target for Opa proteins of Neisseria meningitidis and Neisseria gonorrhoeae. Mol Microbiol *22*, 929-939.

Vitovski, S., Read, R.C.*, et al.* (1999). Invasive isolates of Neisseria meningitidis possess enhanced immunoglobulin A1 protease activity compared to colonizing strains. FASEB J *13*, 331-337.

Vogel, U., Taha, M.K.*, et al.* (2013). Predicted strain coverage of a meningococcal multicomponent vaccine (4CMenB) in Europe: a qualitative and quantitative assessment. Lancet Infect Dis *13*, 416-425.

Walport, M.J. (2001). Complement. First of two parts. N Engl J Med *344*, 1058-1066.

Wandersman, C., and Delepelaire, P. (2012). Haemophore functions revisited. Mol Microbiol *85*, 618-631.

Watanabe, S., Takahashi, N.*, et al.* (2012). Human neuroglobin functions as an oxidative stress-responsive sensor for neuroprotection. J Biol Chem *287*, 30128-30138.

Weinberg, E.D. (1975). Nutritional immunity. Host's attempt to withold iron from microbial invaders. JAMA *231*, 39-41.

Weinberg, E.D. (1978). Iron and infection. Microbiol Rev *42*, 45-66.

Weinberg, E.D. (1997). The Lactobacillus anomaly: total iron abstinence. Perspect Biol Med *40*, 578-583.

Weiss, G. (2005). Modification of iron regulation by the inflammatory response. Best Pract Res Clin Haematol *18*, 183-201.

Wilder-Smith, A., Barkham, T.M.*, et al.* (2002). Acquisition of W135 meningococcal carriage in Hajj pilgrims and transmission to household contacts: prospective study. BMJ *325*, 365-366.

Wilks, K.E., Dunn, K.L.*, et al.* (1998). Periplasmic superoxide dismutase in meningococcal pathogenicity. Infect Immun *66*, 213-217.

Williams, E.W., Hawkey, P.M.*, et al.* (1983). Serious nosocomial infection caused by Morganella morganii and Proteus mirabilis in a cardiac surgery unit. J Clin Microbiol *18*, 5-9.

Wong, S., Lennon, D.*, et al.* (2007). New zealand epidemic strain meningococcal B outer membrane vesicle vaccine in children aged 16-24 months. Pediatr Infect Dis J *26*, 345-350.

Yang, X., Yu, R.H.*, et al.* (2011). Anchor peptide of transferrin-binding protein B is required for interaction with transferrin-binding protein A. J Biol Chem *286*, 45165-45173.

Yazdankhah, S.P., and Caugant, D.A. (2004). Neisseria meningitidis: an overview of the carriage state. J Med Microbiol *53*, 821-832.

Yu, C., and Genco, C.A. (2012). Fur-mediated global regulatory circuits in pathogenic Neisseria species. J Bacteriol *194*, 6372-6381.

Zhao, S., Montanez, G.E.*, et al.* (2010). Regulatory role of the MisR/S two-component system in hemoglobin utilization in Neisseria meningitidis. Infect Immun *78*, 1109-1122.

Zhu, W., Hunt, D.J.*, et al.* (2000a). Use of heme compounds as iron sources by pathogenic neisseriae requires the product of the hemO gene. J Bacteriol *182*, 439-447.

Zhu, W., Wilks, A.*, et al.* (2000b). Degradation of heme in gram-negative bacteria: the product of the hemO gene of Neisseriae is a heme oxygenase. J Bacteriol *182*, 6783-6790.