Summary

M. haemolytica, *H. somni* and *Pasteurella multocida* are the main bacterial agents involved in the multifactorial Bovine Respiratory Disease complex (BRD), along with viral coinfections such as bovine respiratory syncytial virus (bRSV). This disease complex costs the agricultural industry vast sums of money annually due to morbidity, mortality and the costs of metaphylaxis, prophylaxis and treatment. Vaccines and antibiotics are available for certain strains of these species, however, the increasing incidence of antibiotic resistance is decreasing the effectiveness of many of these drugs. It is therefore vital to discover new antimicrobial targets, enabling the continued and ideally improved prevention and treatment of the diseases these pathogens cause.

This project investigated homologues of a protein identified in *Neisseria meningitidis,* known as Gonolysin 1 (Gly1). *N. meningitidis* Gly1 was previously implicated in haem-iron uptake *in vitro,* a phenomenon that was also observed with the *Neisseria gonorrhoeae* homologue. Iron uptake is essential for bacterial survival and is an important virulence factor. A recognised mechanism to limit bacterial growth is iron sequestration; preventing the bacteria from accessing haem or iron sources is therefore a potential target for therapeutic intervention.

Putative Gly1 homologues were discovered in other economically relevant bacterial species. With the aim of characterising these homologues in *M. haemolytica*, *A. pleuropneumoniae* and *Salmonella enterica arizonae (IIIa),* the genes predicted to encode the Gly1 proteins were cloned and overexpressed in *E. coli* cells.

These proteins were studied using a range of experimental techniques to elucidate their biological function. To further characterise the protein structure, biophysical approaches such as circular dichroism, X-ray crystallography and bioinformatics were employed. A crystal structure of the *M. haemolytica* paralogue (MHpara) was deduced, along with the functions of a number of these novel proteins. While the homologues discovered in *M. haemolytica* and *S. enterica arizonae (IIIa)* showed interactions supporting the haem- binding hypothesis, the *A. pleuropneumoniae* homologue and MHpara did not. MHpara did, however, display some adhesion to respiratory tissue. To conclude, not all Gly1 homologues have the same function with some involved in nutrient uptake and some potentially involved in virulence.

Acknowledgements

I would like to thank the AHVLA and The University of Sheffield for funding the project. Thank you to Professor Sayers for giving me the opportunity to undertake this project, and for the input, especially in the last few months. I would also like to thank the members of the Sayers Lab group, Sarbendra, Magda, Roxanne, Jess and Guta who have provided encouragement and entertainment throughout my time here and especially to Janine Phipps who has provided invaluable technical support. A very special mention goes to Sarah Oates and Sam Harding, who have been there through the inevitable ups and downs of the last few years- without their help and friendship I don't think I would have got this far! Thank you also to my friends outside the lab for their belief and understanding whilst I was concentrating on science.

I would like to acknowledge the parts played by members of the Artymiuk laboratory, Jason Wilson and Rick Salmon, for their work on the protein X-ray crystallography and to Yvonne Stevenson for help with histology.

Finally and most importantly I would like to thank my wonderful husband Gareth, my parents and my brother for their love and support over the last four years, it is due to their help and words of encouragement (and occasional provision of meals during long working weeks) that I have got to this point.

List of Figures

List of Tables

Abbreviations

ABC Transporter	-	ATP binding cassette transporter
ACP	-	Adhesin complex protein
AHVLA	-	Animal Health and Veterinary Laboratory Association
APP	-	*Actinobacillus pleuropneumoniae*
APS	-	Ammonium persulphate
ATP	-	Adenosine triphosphate
BHV-1	-	Bovine herpes virus - 1
BLAST	-	Basic local alignment sequence tool
BRD	-	Bovine respiratory disease
bRSV	-	Bovine respiratory syncytial virus
BSA	-	Bovine serum albumin
CBA	-	Colombia blood agar
CD	-	Circular dichroism
CM	-	Cytoplasmic membrane
c.v.	-	Column volume
DNA	-	Deoxyribonucleic acid
dNTP	-	Deoxynucleotide triphosphate
DTT	-	Dithiothreitol
EDTA	-	Ethylenediaminetetraacetic acid
FEN	-	Flap endonuclease
FITC	-	Fluorescein isothiocyanate
Fur	-	Ferric uptake regulator
Gcn5	-	General control of amino-acid synthesis 5
Gly1	-	Gonolysin1
GuHCl	-	Guanidine hydrochloride

hGH	-	Human growth hormone
HmbR	-	Haemoglobin receptor
HO	-	Haem oxygenase
ICE(P)	-	Integrative conjugative element (protein)
IL	-	Interleukin
IPTG	-	Isopropyl-beta-D-thiogalactopyranoside
LBP	-	Lactoferrin binding protein
Lkt	-	Leukotoxin
LOS	-	Lipo-oligosaccharide
LPS	-	Lipopolysaccharide
OMP	-	Outer membrane protein
ORF	-	Open reading frame
PAGE	-	Polyacrylamide gel electrophoresis
PBP	-	Periplasmic binding protein
PBS	-	Phosphate buffered saline
PCR	-	Polymerase chain reaction
RCSB - PDB	-	Research Collaboratory for Structural Bioinformatics - Protein Data Bank
pmf	-	Proton motive force
PMSF	-	Phenylmethylsulfonyl fluoride
PPIX	-	Protoporphyrin IX
PSI-BLAST	-	Position Specific Iterated BLAST
RBS	-	Ribosomal binding site
RNA	-	Ribonucleic acid
ROS	-	Reactive oxygen species
RTX	-	Repeat in toxin

RVC	-	Royal veterinary college
SDS	-	Sodium dodecyl sulphate
TBP	-	Transferrin binding protein
TE	-	Tris-EDTA
TEMED	-	Tetramethylethylenediamine
TMD	-	Transmembrane domain
TNF	-	Tumour necrosis factor
UV	-	Ultraviolet
wt	-	Wild type

1 Introduction

1.1 Key Bacterial Pathogens

1.1.1 Pasteurellaceae

The Pasteurellaceae family of gammaproteobacteria currently consists of 78 species, which are made up of 21 genera; *Mannheimia, Haemophilus, Histophilus* and *Actinobacillus* form the genera of interest in this family (Euzeby, 1997). They are all Gram-negative, rod shaped, non-motile facultative anaerobes. These species have been reclassified a number of times in recent decades due to the development of more sophisticated genetic identification techniques. Within the Pasteurellaceae family are species that are capable of existing as commensals and/ or pathogens in vertebrates. *Haemophilus influenzae,* for example, is an opportunistic human pathogen, often found as a commensal yet able to cause respiratory disease. The various species of Pasteurellaceae can be differentiated by their phenotypes such as haemolytic activity, indole production, β-NAD dependence and enzymatic activity as well as analysis of 16S sequences (Dousse et al., 2008).

Other bacteria from this family inhabit similar niches in many other species, colonising the mucosal surfaces of the respiratory tract, oropharynx and reproductive tracts. For example, *Mannheimia haemolytica* exists as a commensal in ruminants, mainly found in the bovine and ovine nasopharynx. These bacteria are again able to make the switch from commensal to pathogen when the host is stressed or immunocompromised, leading to respiratory disease. Infection often occurs during transport of herds of animals- known as 'feedlots'- between farms giving rise to the name 'Shipping Fever'. This is due to the heightened stress caused by transportation and the close proximity of the animals being moved. In the case of cows this is known as Bovine Respiratory Disease (BRD).

In addition to BRD, this bacterial family is responsible for infections in many mammals and other vertebrates from cows, to pigs, to chickens. The major pathogens isolated from diseased animals are *Pasteurella multocida, Bibersteinia trehalosi, M. haemolytica, A. pleuropneumoniae, Haemophilus parasuis,* and *Avibacterium paragallinarium*

respectively. Disease in these species is responsible for huge economic losses in the agricultural industry. Pneumonia is the most common symptom of disease arising from Pasteurellaceae infections, although further pathology does occur including mastitis in sheep and severe systemic infection.

In addition to the respiratory tract, *H. somni* also resides in the reproductive mucosa and, like other bacterial species mentioned, can cause systemic disease in the bovine host with comorbidities including septicaemia, infertility and abortion, myocarditis and meningoencephalitis (Harris et al 1989). These systemic infections enhance the economic burden of the bacteria causing BRD to the agricultural industry. *A. pleuropneumoniae* is a commensal of the porcine host, where it again can be isolated from asymptomatic carriers, generally found inhabiting the tonsillar crypts and nasopharynx (Bisgaard, 1993). This too is able to cause disease; porcine pleuropneumonia can be highly contagious, with high mortality rates depending on the strain causing the outbreak (Bandara et al., 2003). Similar to other members of the *pasteurellaceae* family it causes severe necrotic lung lesions (Ajito et al., 1995). It can also lead to pericarditis and abortion in pigs, costing the US economy vast sums annually, reported to be in excess of $32 million in 1995 (Losinger, 2005).

1.1.2 Salmonella

Salmonella species are members of the *Enterobacteriaceae* family. There are two species- *S. bongori* and *S. enterica*, which are split into 12 subspecies and further subdivided into over 2000 serovars according to the NCBI database. They are Gram-negative, rod shaped facultative anaerobes that can cause infections in a wide range of species. The pathology resulting from these bacterial infections depends upon both the serovar and the host. The non-typhoidal *S. enteretidis* causes gastroenteritis in humans, *S. typhi* can cause typhoid fever in many hosts depending on the bacterial sub-species and *S. enterica arizonae (IIIa)* resides as a commensal in its cold-blooded host. Both humans and a wide variety of animals are carriers or reservoirs for *Salmonella*, and the bacteria can be food-borne or spread by contact with an infected party. This ability to survive and multiply in numerous different host species enhances the potential for disease outbreaks.

The typhoidal *S. enterica* strains often begin as a gastric infection, being taken into the epithelium by pinocytosis after ingestion. The infections can then become systemic, invading by crossing the intestine using the host immune defence to its advantage to enter the blood stream. Salmonella infections represent a huge public health issue that is exacerbated by the rise in antibiotic resistance, limiting the treatment options available. In Asia, typhoid caused by *S. typhi* is considered endemic, partly due to lack of vaccination and overcrowding in certain locations. Salmonella outbreaks are therefore still a concern in the Western world. The CDC recently reported that non-typhoidal Salmonella infections are one of the leading causes of food borne hospitalisation in the USA (Scallan et al., 2011).

S. enterica arizonae (IIIa) is rarely isolated in cases of human disease. It is, however, a common commensal of reptiles and amphibians, estimated to contribute to 8.3 % of the *Salmonella* species colonising wild reptiles (Briones et al., 2004). Contact with these species is the main cause of disease in humans caused by *S. arizonae*. Some cases of transmission to poultry have also been reported, highlighting the danger of enzootic transmission. Reptile acquired Salmonella infection in general accounts for approximately 6% of human salmonellosis overall (Mermin et al., 2004). Infection mainly manifests as gastroenteritis, although cases of meningitis (Lakew et al., 2013) and osteomyelitis (Kolker et al., 2012) have also been reported.

1.1.3 Neisseriaceae

The family *Neisseriaceae* are Gram- negative, non-motile, diplococcoid, capsular bacteria. It contains 5 main genera; *Acinetobacter, Kingella, Moraxella*, *Oligella* and *Neisseria*. The genus *Neisseria* contains species of bacteria known to be the main causes of disease in humans from this bacterial family. *Neisseria* species are found within the human host with some inhabiting niches in other mammals such as dogs and monkeys. The majority of these species are commensals of the respiratory and reproductive mucosa, such as *N. cinerea*. Some, such as *N. lactamica*, while often isolated from the nasopharynx of asymptomatic carriers, are rarely associated with disease in immunocompromised hosts. The most important human pathogens of this genus are *N. meningitidis* and *N. gonorrhoeae*.

N. meningitidis is the major global cause of meningococcal meningitis and infection can lead to septicaemia and death. While 10% of the population are asymptomatic carriers of the bacteria, when disease occurs the mortality rate from meningococcal meningitis is between 10- 15% in the US. Of the survivors, 11- 19% will suffer serious long term effects (CDC, 2015). These devastating sequale include loss of vision, deafness and loss of limbs (Heckenberg et al., 2008). There are 12 serogroups of *N. meningitidis* which are differentiated according to their capsular polysaccharides. Five of these serogroups, A, B, C, W and Y, are commonly implicated in causing disease globally (Harrison et al., 2009). Cases of meningococcal meningitis have decreased in many countries over recent years due to the availability of vaccines to serogroups A, C, W and Y. However, the efficacy of the available vaccines to different serotypes is variable, which led to the multivalent vaccine MenACWY-CRM. These are based on the polysaccharide components of the bacterial serotypes. Serotype B vaccines were unable to be developed using this technique due to the similarity of the polysaccharide to a component of human cells. A recent review summarised the development of the 4CMENB vaccine against *N. meningitidis* serotype B (Andrews and Pollard, 2014).

N. gonorrhoeae is also an important human pathogen, causing the sexually transmitted infection gonorrhoea. Gonorrhoea infects both men and women. It is characterised by urethritis and cervicitis, leading to pelvic inflammatory disease, infertility and ectopic pregnancy if not resolved. The global burden of *Neisseria gonorrhoeae* infections is exacerbated by rising antimicrobial resistance. Globally, approximately 106 million of the curable sexually transmitted infections are due to *N. gonorrhoeae* each year which equates to around 20% (WHO, 2012).

1.2 Bovine Respiratory Disease

1.2.1 Causes and Importance

The Bovine Respiratory Disease (BRD) complex, as the name suggests, is a multifactorial disease. It is caused by the interplay between viral and bacterial agents and the host immune system. *Mannheimia haemolytica* and *Pasteurella multocida* are the most commonly isolated bacterial agents, with *H. somni* also less commonly isolated from diseased cattle. These bacteria also infect sheep and *B. trehalosi* is an additional factor found in isolated in ovine disease cases. Bovine respiratory syncytial virus (bRSV), bovine viral diarrhoea virus and bovine parainfluenza virus are all among common concomitant viral infections involved in the overall pathogenesis of the disease. *M. haemolytica* is also a well-documented cause of ovine mastitis, which has a detrimental effect on the sheep farming industry.

BRD presents a huge challenge to the agricultural industry. The morbidity and mortality rates in the US alone cost the economy $1 billion annually (USDA, 2011), with costs of prevention and treatment also increasing year-on-year. New prevention or treatment measures are desperately required as antibiotic resistance increases.

1.2.2 Host- Pathogen Interactions

1.2.2.1 Mechanism of Colonisation and Infection

M. haemolytica is the bacterial species most often isolated from diseased animals. It is also known to be a commensal of the bovine upper respiratory tract and nasopharynx, yet the mechanism of the switch from commensal to pathogen is not fully understood. A study into the bacteria commonly isolated in North America and Canada found that serotypes A1, A2 and A6 are prevalent. A1 and A6 are the serotypes generally isolated from diseased cattle (70 and 20%, respectively), with a lesser percentage of A2 (10 %). The study also showed that A2 was the predominant serotype isolated from healthy cattle (Klima et al., 2014a). Serotype A2 is found in other ruminants and is the primary serotype indicated as causing ovine respiratory disease. The current hypothesis is that

the A1 and A6 strains outgrow the A2 serotype as a result of additional factors such as co-infection, leading to disease.

Bovine respiratory disease is characterised by initial symptoms of fever, nasal discharge, cough, lack of appetite and anorexia, which in itself affects the value of the cattle due to the quality and quantity of meat, before considering that approximately 5.9% of cattle with BRD die from the infection (Faber et al., 1999). A severe infection can lead to fibronecrotising pleuropneumonia. The majority of factors enabling colonisation are also virulence factors, which work against the host defence systems to promote disease. These processes as a whole, or their individual protein components, therefore present potential targets for prophylaxis or treatment.

1.2.2.2 Virulence Factors

In the case of *M. haemolytica,* the bacteria- either already present as a commensal or transmitted from infected animals- grows rapidly in the host upper respiratory tract where it is located in the mucous layer. This mucous is secreted by either goblet cells or seromucous glands depending on the location within the respiratory tract. It contains antibacterial agents such as lysozyme, secreted antibodies and host antimicrobial peptides including β-defensins, which are innate defence mechanisms that attack the peptidoglycan walls of the invading bacterial cells (Ackerman et al., 2011). Another innate host defence mechanism is the mucociliary apparatus allowing clearance of pathogens captured in the mucous barrier. However, this is not always successful at clearing the infection, with bacteria eventually getting dislodged, inhaled, and settling in the lower respiratory tract or lung (Grey and Thomson, 1971).

In healthy animals, the immune system is able to clear this infection- when the bacteria enter the lung, alveolar macrophages are activated by lipopolysaccharide (LPS) on the surface of *M. haemolytica*, causing the release of pro-inflammatory cytokines. These cytokines (e.g. IL-1, TNF-α and IL-8) signal to the host immune system and aid in the recruitment and priming of neutrophils (Hsuan et al., 1999). Neutrophil recruitment causes the release of further cytokines, reactive oxygen species and elastase which act to damage bacterial cell membranes. Expression of L- selectin on the neutrophil surface enables the cells to adhere to β2-integrins on the host respiratory epithelium allowing

the neutrophils to cross into the tissue to clear the infection. Phagocytosis by both neutrophils and tissue macrophages clears the bacteria and any cellular debris. Mast cells, lymphocytes and airway epithelial cells also play a vital part in responding to the infection.

In stressed, immunocompromised animals that may have a viral co-infection, these defence mechanisms can fail and colonisation can lead to severe respiratory disease. Leukotoxin and LPS are the major virulence factors of *M. haemolytica, H. somni* and *A. pleuropneumoniae* playing key roles in disease severity. LPS can act as a protective barrier to the bacterial cell, preventing certain antimicrobial peptides from crossing the outer membrane and damaging or killing the bacterium. Descent of bacteria to the lower respiratory tract and lung leads to cellular infiltration of the alveoli by leukocytes causing a 'pro-coagulation' state. This results in fibrin polymerisation and necrosis of host tissues due to the cytokines and reactive oxygen species released (Car et al., 1991). Leukotoxin stimulates the release of leukotriene B4, histamine and prostaglandin E2, all of which contribute to the severe inflammatory damage in the infected lung (Adusu et al., 1994). At higher concentrations Lkt can also induce apoptosis in host defence cells, impairing the response to disease. A leukotoxin knock-out strain of *Mannheimia haemolytica* was created by Tatum *et al* which showed a dramatic decrease in host cellular damage in a bovine model of disease after infection with the knock-out strain (Tatum et al., 1998).

Lafleur *et al* demonstrated an increased pathogenicity driven by LPS and Lkt together, as evidenced by an increase in cytolysis of alveolar macrophages due to Lkt in the presence of non-cytolytic levels of LPS, compared to Lkt alone (Lafleur et al., 2001). Prior exposure to Lkt or LPS increased the production of β-defensins which actually enhanced the cytotoxic activity (Leite et al., 2002). The leukotoxins of *Mannheimia haemolytica* have been shown to vary depending on the serotype, with slightly differing levels of toxicity to bovine and ovine neutrophils (Davies and Baillie, 2003).

A. pleuropneumoniae infection in pigs causes necrotic lung lesions in a similar way, leading to the production of pro-inflammatory cytokines and the subsequent characteristic accumulation of neutrophils. Viral co-infection is a predisposing factor to BRD. BHV-1 has been shown to increase the pathogenicity of *M. haemolytica* by

increasing the amount of LFA-1 (a β2-integrin) produced by host cells, which is known to be involved in Lkt binding (Leite et al., 2002, Rivera-Rivas et al., 2009).

Endotoxins produced by *H. somni* and *A. pleuropneumoniae* are also implicated in bacterial virulence. Lkt is a member of the RTX family of toxins, to which the major virulence factor proteins of *A. pleuropneumoniae*, Apx, also belong (Jansen et al., 1995). This family of toxins target both macrophages and neutrophils creating pores through the membranes which leads to cytolysis (Thumbikat et al., 2003) and, via release of reactive oxygen species (ROS), contributes to the damage of the pulmonary epithelium. *H. somni* also possesses the ability to damage bovine endothelial cells via lipo-oligosaccharide (LOS) with a similar pore forming mechanism (Sylte et al., 2001). Rtx toxins are also known hemolysins, potentially providing a cellular mechanism for nutrient acquisition in low iron environments, although there is little evidence to support this.

The bacterial polysaccharide capsule prevents phagocytosis by the host immune system, as demonstrated by a capsule negative strain of *M. haemolytica* that was both phagocytosed more readily than the wild type bacteria and more susceptible to complement-mediated killing (McKerral and Lo, 2002, Chae et al., 1990). The capsule can also interact with proteins in the pulmonary surfactant, increasing bacterial adhesion. *M. haemolytica* serotypes have been demonstrated to have differing capsular polysaccharide composition, which may be relevant to the virulence of each serotype (Adlam et al., 1984, Adlam et al., 1986, Puente-Polledo et al., 1998). Cellular adhesion is considered another important virulence factor, enabling the initial host colonisation and allowing, if the conditions are appropriate, invasion of the host.

Outer membrane proteins (OMPs) are also involved in adhesion when not masked by the capsule, as seen in many bacterial species, such as OmpA in *M. haemolytica* (Hounsome et al., 2011) and OMPs of *A. pleuropneumoniae (Van Overbeke et al., 2002)*. Of the mannheimia outer membrane proteins, OmpA and PlpE are known to be major targets of host antibodies. Outer membrane proteins have been demonstrated to be vital in bacterial host colonisation, especially iron regulated OMPs, and it is therefore understandable that they present a target for the host immune system. OmpA1 is found

in bacterial isolates from the bovine host whereas OmpA2 is found in strains of ovine origin (Hounsome et al., 2011), which demonstrates some species specificity. These and other OmpA proteins differ between serotypes in the hypervariable extracellular loop region (Hounsome et al., 2011). OM components of *M. haemolytica* are outlined in Figure 1.1.

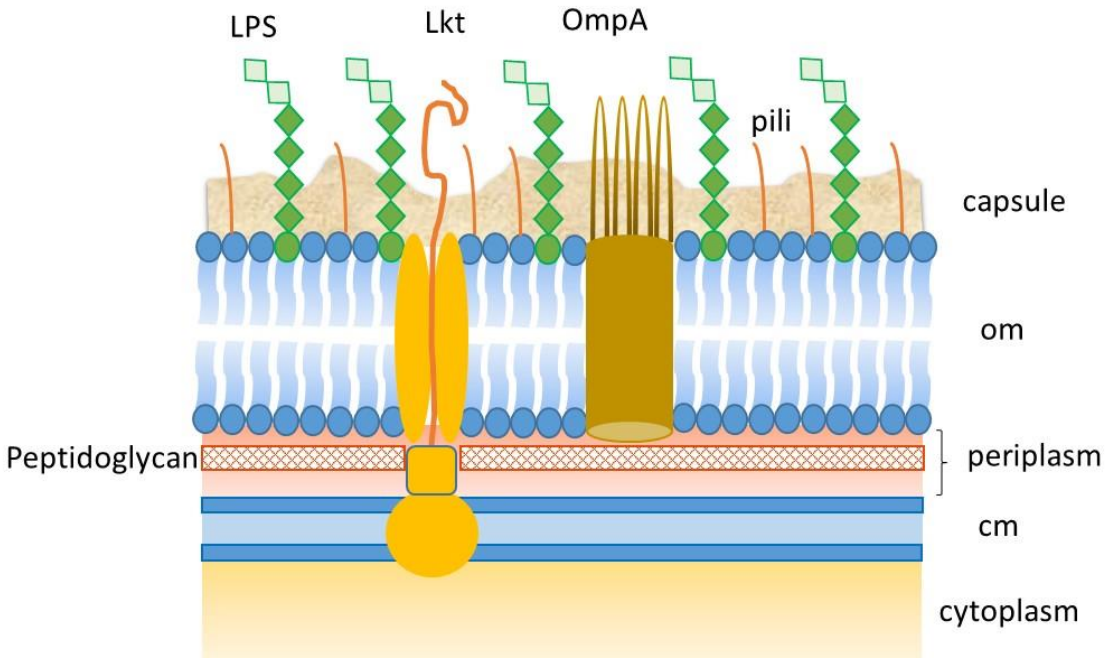


Figure 1.1 Outer Membrane Virulence Factors of *M. haemolytica*.

LPS (green) with the O antigen (depicted in light green), pili on the cell surface (orange) and Lkt secreted from the cell via a Type I secretion system comprised of LktB and LktD in *M. haemolytica* (yellow). OmpA shown with characteristic 4 extracellular loops and a membrane bound beta barrel (brown). The capsule of *M. haemolytica* is shown on the extracellular face. Om - outer membrane, cm - cytoplasmic membrane. Adapted from Gill EE, Brinkman FS - Bioessays (2011), open access.

Virulence factors of *M. haemolytica* differ between serotypes and host species. Lkt is found throughout *M. haemolytica* serotypes, as is LPS. While Lkt is relatively well conserved, LPS of *M. haemolytica* differs between serotypes in both the core region and O antigen (Davies and Donachie, 1996). LPS of *A. pleuropneumoniae* is also involved in adhesion along with the fimbriae and a number of other OMPs (Van Overbeke et al.,

2002). Some LOS in *H. somni* and other bacterial species have been found to mimic mammalian LOS in an attempt to impair immune surveillance (Mandrell et al., 1992). LPS has been shown to enhance the effect of the Apx toxins of *A. pleuropneumoniae* on phagocytes and further activate the host inflammatory response in a study comparing the wild type strain to a LPS mutant strain (Ramjeet et al., 2008).

Alterations in the components of the lipo-oligosaccharide component, such as phase variation and sialylation, have been identified as virulence factors in *H. somni*, enabling the bacteria to evade the host immune system (Inzana et al., 1992). Other outer membrane proteins involved in virulence include receptors for host iron-containing proteins such as transferrin and lactoferrin. Many proteins involved in nutrient uptake such as haem and iron receptors are iron regulated, demonstrating the importance of iron uptake for bacterial survival and colonisation. An important factor in colonisation and pathogenesis is the presence of essential nutrients, and the ability of the bacteria to capitalise on what is available. The host employs methods to reduce the 'free' nutrients available to limit the possibility of colonisation and to protect itself from cellular damage in the case of iron and other reactive species.

1.3 Iron and Haem Acquisition

Bacteria such as *M. haemolytica* reside within the upper respiratory tract of ruminants. The availability of iron within these mucosal surfaces is low compared to the blood, yet iron uptake is essential for bacterial survival and is therefore counted as an important virulence factor. Iron is used in numerous cellular processes from cellular respiration, to regulation of transcription, to DNA repair. Different bacteria obtain this iron from a variety of sources depending upon the ecological niche they inhabit. The host environment generally maintains a low level of free iron by sequestering it from circulation in high affinity complexes such as lactoferrin and transferrin. These proteins bind free iron both to prevent toxicity to the hosts own cells and to reduce nutrients available to invading organisms.

Iron exists in equilibrium between ferrous (Fe^{2+}) and ferric (Fe^{3+}) forms, depending on the surrounding environment but is generally found in its insoluble ferric form within living things i.e. at a neutral pH (Fontecave and Pierre, 1993). Iron is a core component of haem, which itself is integral to life in the form of diverse hemoproteins such as haemoglobin and cytochrome C. Too high a concentration of iron or haem can lead to cellular toxicity due to the formation of reactive oxygen species (ROS) and free radicals in the Fenton Reaction ($Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + {}^{\cdot}OH + OH^-$) causing DNA and protein damage. It is for this reason that haem levels are tightly controlled within the host. With the majority of iron being complexed in haem or hemoproteins in vertebrate systems, these are important targets for pathogens. Haem is often found in the ferrous form within cells, whereas hemin, the oxidisation product of haem, generally exists in extracellular niches and is therefore the form involved in the uptake process as part of hemoproteins.

Bacteria have evolved many different ways of surviving in environments where nutrients are scarce, which include secretion of molecules that chelate iron with a high affinity and proteins to retrieve haem known as siderophores and hemophores, respectively as summarised in Figure 1.2. These scavenger systems retrieve nutrients for the bacterial cell, binding to specific outer membrane receptors for internalisation. Periplasmic binding proteins then transport the nutrient to the cytoplasmic membrane ATP binding

cassette (ABC) transporters, which shuttle it into the cytoplasm where it is used or stored. All of these systems are tightly regulated due to the reactive nature of haem.

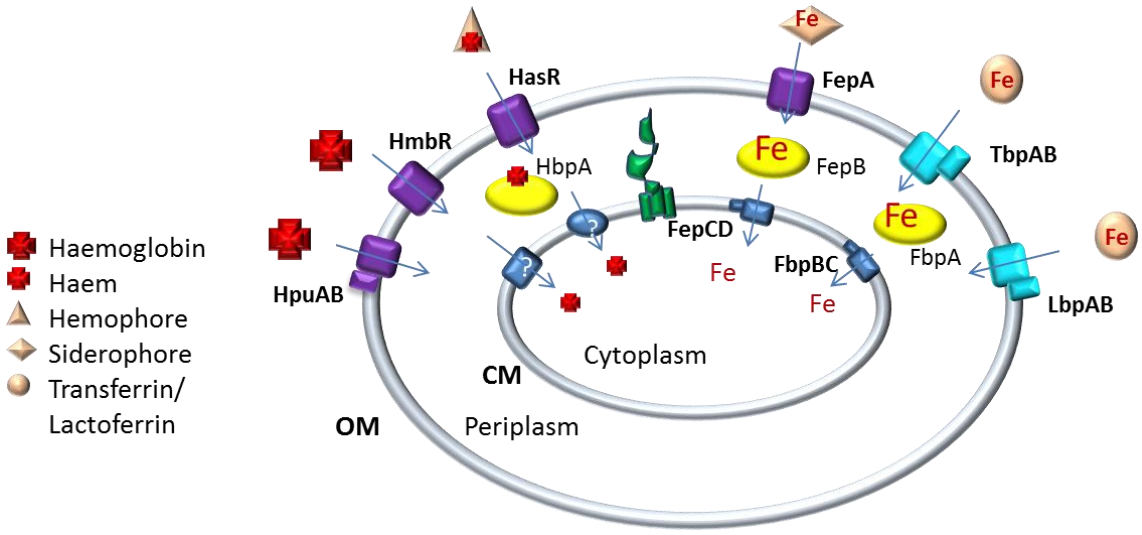


Figure 1.2 Haem and Iron Uptake Systems in Gram-negative Bacteria.

Gram-negative bacteria produce specific outer membrane proteins (purple and blue) to capture iron/ haem protein complexes. The substrate binds to the OM receptor and the iron/ haem is internalised using the pmf delivered by the TonB complex (green). The periplasmic binding proteins (yellow) shuttle the internalised molecule across the periplasm to an ABC transporter (navy blue) where it is transported into the cytoplasm for use or storage. Adapted with permissions (Anzaldi and Skaar, 2010).

1.3.1 Siderophores and Hemophores

Siderophores and hemophores are secreted by certain bacterial species to scavenge iron and haem from the surrounding environment. They bind these molecules with moderately high affinity, for example HasA of *Serratia marcescens* binds hemin with a K_d of 5.3 x 10^{-10} M (Deniau et al., 2003), and delivers it to the bacterial cell. The nutrient uptake occurs by the binding of the secreted chelating agent to specific outer membrane proteins. Hemophores such as HmuY from *Porphyromonas gingivalis* have been shown to have a higher affinity for hemin than haem whilst retaining the ability to bind the reduced form (Smalley et al., 2011). The mechanism of initial haem removal from haemoglobin is hypothesised to be by passive diffusion based upon the higher affinity of hemophores for haem than haemoglobin for example (Letoffe et al., 1999). A novel proposed method of haem capture centres on the enzymatic oxidation of the ferrous

iron at the centre of the haem molecule. This decreases the affinity of haemoglobin for haem, allowing HmuY to scavenge the haem (Smalley et al., 2011).

Siderophores of *E. coli,* for example, include ferric enterobactin, which binds to the receptor FepA, and ferrichrome, which binds to the OM protein FhuA. The mechanism of siderophore- OM receptor interaction has been described by Locher *et al*. Ferrichrome binds, iron-complexed region first, to an aromatic pocket located on the cell surface side of the β-barrel, which in its closed form is obstructed by a plug. The ferrichrome binding site comprises loops L3 and L11 of the β-barrel and part of this plug. Binding to the receptor leads to a conformational change whereby part of the plug region shifts towards the ligand, causing the backbone to be displaced and a subsequent N-terminal conformational alteration (Locher et al., 1998). This movement is hypothesised to be the signal to the TonB receptor to produce the energy required to bring the iron into the cell. *N. gonorrhoeae* is an example of a bacterial species that do not produce their own siderophores. The bacteria compensate for this by utilising those secreted by other cells, or 'xenosiderophores', internalising the iron from enterobactin via the fbpABC proteins in a TonB independent manner (Strange et al., 2011).

The hemophore HasA, secreted by *Serratia marcescens*, was the first to have its structure solved and has therefore been widely investigated. It consists of an α/β protein fold, and the haem molecule is held between two loops within this; L1 and L2. Conformational alterations in this site are key in the binding and release of the haem molecule (Wolff et al., 2008). As mentioned, hemophores bind haem with very high affinity; the iron of the haem binds to HasA via hydrogen bonding to His-32 and Tyr-75 (Caillet-Saguy et al., 2012), and delivers the haem to a specific OM receptor, HasR (Figure 1.3). They are able to acquire both free and complexed haem. A hemophore found in pasteurellaceae species, HxuA, is able to acquire haem from hemopexin (Gioia et al., 2006). The mechanism of haem acquisition by hemophores has been described as either competition with the hemoprotein to appropriate the haem or by causing conformational changes to liberate the haem from its carrier protein, thereby allowing the hemophore to capture the haem.

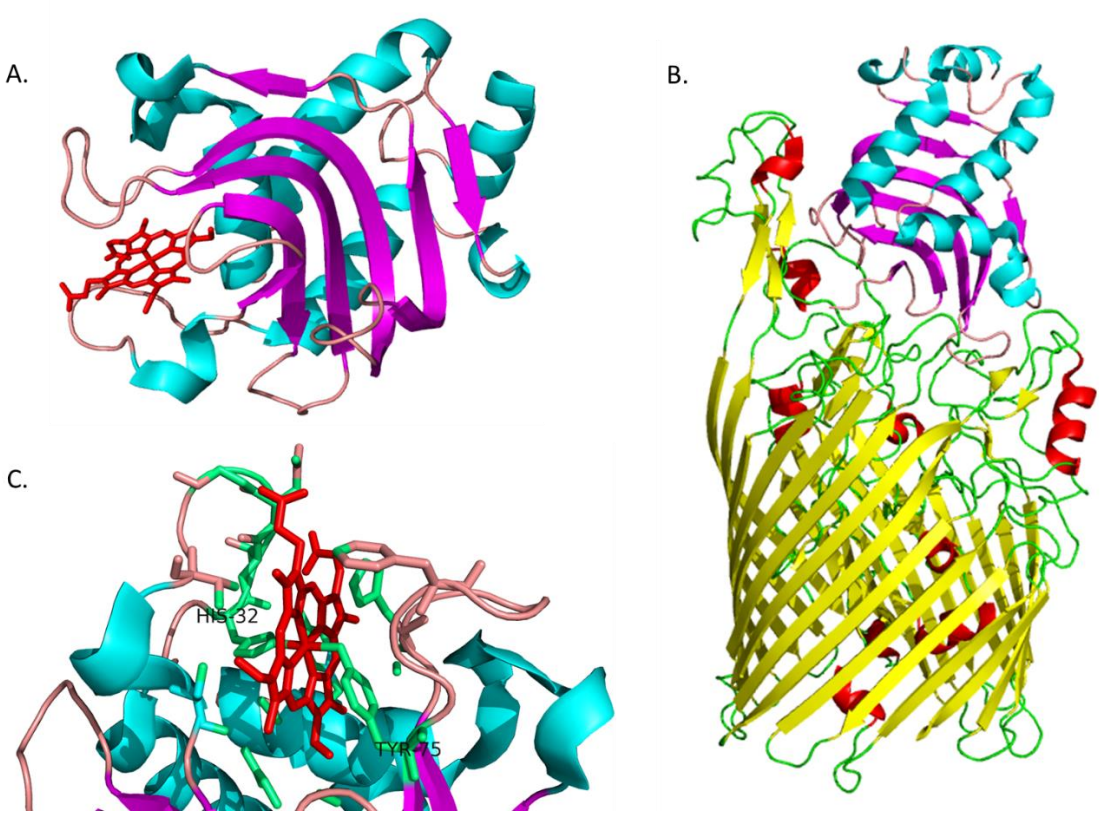


Figure 1.3 Hemophore Structures and Interactions with Haem.

A. shows the HasA monomer crystallised with bound haem in red sandwiched between the loops (PDB code 1DK0), **C.** shows the haem coordinating amino acids- His-32 and Tyr-75 and **B.** shows the hemophore bound to HasR (PDB code 3CSN). Structures were both *S. marcescens* taken from RSCB-PDB Protein workshop and ray traced and rendered in Pymol.

HasR, the HasA receptor, binds holo-HasA between loops L6 and L9, with the haem molecule facing the receptor. The haem is bound by two histidines within the receptor, His-189 and His-603 on the plug and barrel respectively (Izadi-Pruneyre et al., 2006). HasR has a slightly weaker affinity (*Ka* = 5 x 10^{-9} M) for haem than HasA (*Ka* = 5.3 x 10^{-10} M). The transfer of haem between the two proteins is thought to be driven by protein-protein interactions that rely on both the aforementioned histidines, and the conformational changes caused by the hemophore interacting with the receptor. The HasA: HasR binding was reported with an affinity of >10^{-9} M, and this is proposed to be the method of overcoming the minimal difference in affinity for haem between the two proteins (Izadi-Pruneyre et al., 2006). This differs from other haem and iron uptake

systems which utilise proton motive force (pmf) to internalise the nutrient as outlined in section 1.3.3. In this case the pmf is required only for haem internalisation (Izadi-Pruneyre et al., 2006). Examples of this acquisition method have been demonstrated in *P. aeruginosa, S. pyogenes and S. marcesens*. Once the haem has been transferred to the receptor, the apo-HasA dissociates enabling the haem to be transported into the bacterial cell. The haem is then internalised using proton motive force generated by a proton gradient from the cytoplasm to the outer membrane of the cell. This electron transport chain leads to a positive outer membrane and a negative internal environment, resulting in alterations to OM proteins.

1.3.2 Bacterial Outer Membrane Receptors

Haem is a lipophilic molecule and alone is unable to cross the negatively charged outer membrane, therefore transmembrane proteins are required to allow haem uptake into the cell. Known OM receptors involved in haem and haemoglobin uptake include, but are not limited to, HpuAB and HgbA found in neisserial species and pasteurellaceae species such as *H. somni,* and the previously mentioned haem/ hemophore receptor HasR from *S. marcescens*.

HmbR is one outer membrane protein involved in haem uptake from haemoglobin found in neisserial species. A similar as yet uncharacterised potential homologue of HmbR has been discovered in *M. haemolytica* along with further putative homologues in a number of other Gram-negative bacteria. These outer membrane receptors are predicted to have a general basic structure of a beta barrel creating a pore through the membrane, which is blocked by an N-terminal plug. Based on structural modelling, the meningococcal HmbR is predicted to consist of 22 beta strands, 11 surface exposed loops, 10 periplasmic loops and an N-terminal plug (Evans et al., 2010). The haem is again bound to the protein by histidine and tyrosine residues and exists with ferric iron at its centre as a 5-coordinate high-spin molecule (Mokry et al., 2014). No crystal structure of this neisserial HmbR protein has been published to date, however the ShuA haem/ haemoglobin receptor (PDB code 3FHH) bears some similarities to that described by Evans *et al* (Evans et al., 2010). It is mostly composed of beta strands, and the structure depicts 22 of these; the classic number of beta strands forming the barrel structure of haemoglobin receptors. This basic structure was also observed in the Phyre2 predicted tertiary structure of the *M. haemolytica* haemoglobin receptor when analysed with Pymol (Figure 1.4).

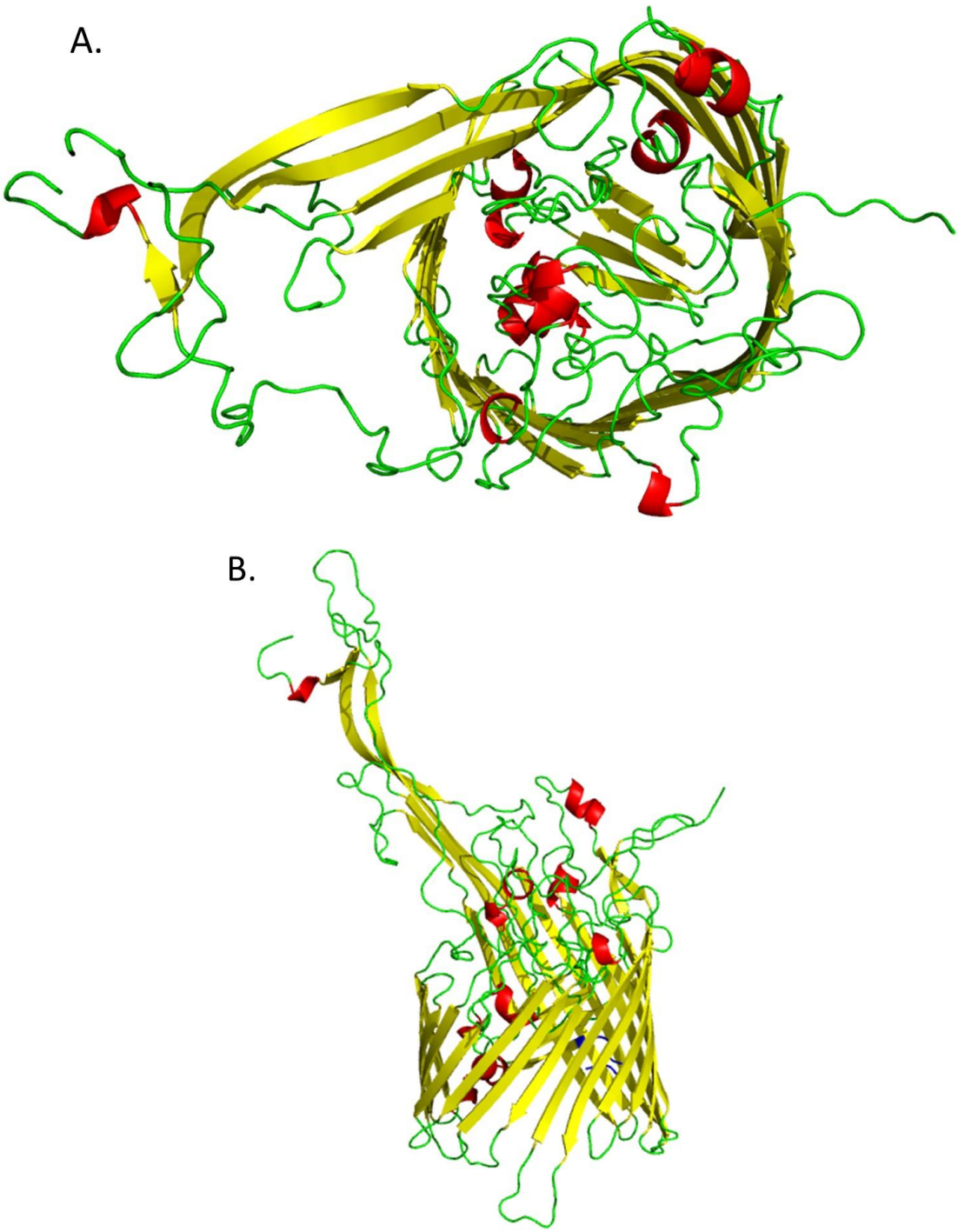


Figure 1.4 Predicted Structure of the *M. haemolytica* HmbR.

Predicted structure of putative *M. haemolytica* HmbR (accession number EDN75148) generated with Phyre2 and rendered and ray traced in Pymol. The template protein used was the crystal structure of TbpA from *Neisseria meningitidis* serogroup b in complex with the c-lobe of human transferrin (3V89). **A.** shows the plug within the beta barrel structure, and **B.** shows the structure from the side with the N- terminal highlighted in blue at the lower right hand side of the plug within the barrel. The predicted structure had 93% coverage (668 amino acid residues) and 100 % confidence.

The neisserial HmbR is a TonB dependent receptor (see section 1.3.3) and is predicted to bind haem via surface loops 2 and 3, with loops 6 and 7 hypothesised to be involved in haem utilisation, possibly to bring haem into the periplasm (Evans et al., 2010). Haemoglobin receptors lacking the plug domain are unable to acquire haem from haemoglobin, suggesting that this component is involved in binding or transport (Fusco et al., 2013). Haemoglobin binding by HgbA of *H. ducreyi* has been shown by mutagenesis studies to occur partly via a series of residues in loop 5 along with histidines in the plug region, whereas the phenylalanine of the 'FRAP' motif on loop 7 was found to be vital in haem uptake (Fusco et al., 2013). Srikumar *et al* identified the protein HgbA in *Actinobacillus pleuropneumoniae* (APP). This appears to form a part of a whole functional haemoglobin receptor, as although a *hgbA* mutant was able to bind haemoglobin, the bacteria was unable to utilise it as a source of iron suggesting that it is a multicomponent receptor complex (Srikumar et al., 2004).

Two-component receptors whose function is to bind lactoferrin or transferrin were initially discovered in bacteria, followed by the identification of a two-component haemoglobin receptor in *N. meningitidis*, HpuAB. These receptors generally consist of a surface exposed lipoprotein, which interacts with the iron-containing substrate, and a transmembrane protein enabling transport into the cell. The neisserial transferrin receptor 'transferrin binding protein AB' (TbpAB) proteins were not found to co-localise on the bacterial membrane, with TbpB forming a surface exposed lipoprotein that binds iron laden-transferrin and TbpA a membrane spanning protein that interacts with TonB (Powell et al., 1998). This TonB interaction allows the protein to transfer the energy required to internalise the substrate (see Section 1.3.3). While the majority of two component receptors require both proteins to be functional, TbpA of TbpAB can act alone as it is capable of binding transferrin and extracting the iron from within, however both components are required for optimal uptake (Stokes et al., 2005).

Lactoferrin binding protein (Lbp) also works as a dual component receptor, as displayed in Figure 1.2, consisting of two OM proteins that are co-transcribed. The two proteins in *N. meningitidis* are again an outer membrane lactoferrin-binding lipoprotein LbpB and the transmembrane protein LbpA. LbpA uses energy provided by the TonB protein to internalise the iron in a similar manner to TbpAB. These proteins are well conserved and

the neisserial LbpB protein has a 55% amino acid sequence identity with TbpB (Lewis et al., 1998).

1.3.3 Cellular Uptake Mechanisms

The transport of proteins across the outer membrane of the bacterial cell requires both an OM receptor and energy to assist the uptake process, however, the OM is not equipped to provide said energy alone. The uptake is powered by a proton motive force (pmf), which exploits the ion gradient throughout the cell. HasR, as previously mentioned, uses the pmf to bring haem across the membrane (Izadi-Pruneyre et al., 2006). The pmf is transmitted from the cytoplasm to the outer membrane via the TonB complex, which consists of three protein subunits; TonB, ExbB and ExbD (Figure 1.5). These proteins are known to form homo-multimers however the stoichiometry has not yet been deduced. ExbB and ExbD are transmembrane proteins located anchored in the cytoplasmic membrane. They facilitate the transfer of the pmf from the cytoplasm outwards, with ExbB transmitting it to ExbD and on to TonB. The N-terminal of TonB also sits in the cytoplasmic membrane.

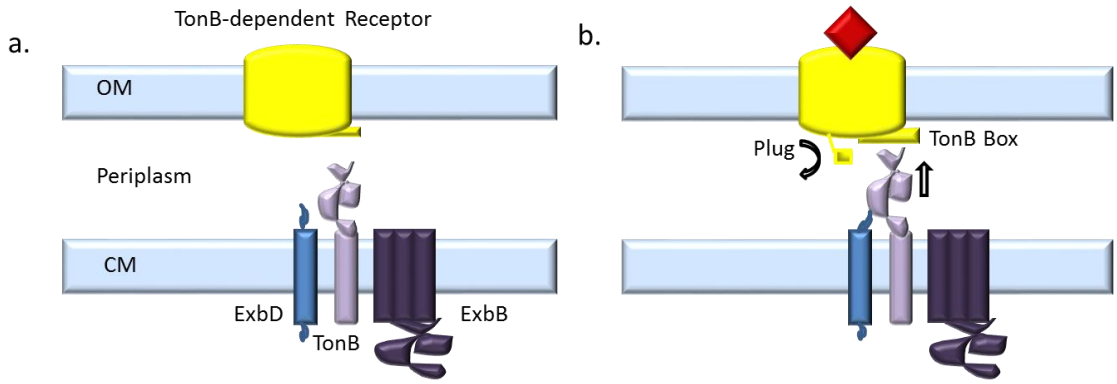


Figure 1.5 TonB-ExbB-ExbD Complex.

Panel **A.** shows the uptake system at rest with the TonB protein (lilac) separate from the TonB Box of the outer membrane receptor, and the plug still in place. **B.** Upon the OM receptor binding a substrate, the pmf allows the TonB C-terminal to contact the TonB Box on the N-terminal of the OM receptor allowing uptake of the substrate. This is a representation of the proteins involved in this process and does not accurately display the ratios of each protein within the TonB complex and cytoplasmic membrane.

The pmf is also required for the cross linking of ExbB and TonB, and can only be transmitted to the OM protein if all three proteins within the complex are intact. An early theory of how TonB transmitted the required 'energy' was the 'Shuttle Hypothesis' which suggested that the TonB protein detached from the cytoplasmic membrane and shuttled across the periplasm to the OM receptor in an 'energised' state to transmit the energy. This theory was replaced after a study by (Gresock et al., 2011) showed the findings to be due to non-specific binding of the dye used in the experimental procedure to other proteins. The method of energy transmission is now believed to involve pmf transfer via the C-terminal of TonB, which contains a proline-rich region able to span the periplasmic space. This contacts both the ExbB-ExbD complex via the C-terminal of ExbB on the cytoplasmic membrane (which is also integral in the association of ExbD and TonB), and the TonB Box. The TonB Box is on the N-terminal region of the outer membrane receptor which contains a motif rich in asparagine and valine residues over 6 or 7 amino acids. TonB-TonB Box contact allows the pmf to transfer to the OM protein and enables a conformational change in the receptor- although the exact mechanism of pmf transmission by TonB is not yet fully understood (Jana et al., 2011).

The pmf provides energy to both remove haem from the holo-protein and to reversibly dislodge the β-barrel plug facilitating the passage of haem through the pore into the cell (Flores Jimenez and Cafiso, 2012). Conformational changes allow all three proteins in the TonB complex to be recycled after the pmf has been transmitted to the OM receptor. TonB proteins are not limited to their role in iron and haem uptake, but are involved in the acquisition of many other nutrients- it is no surprise in this case that other TonB homologues and paralogues have been discovered, such as the *S. marcesens* HasB, which while the overall function is conserved, do not necessarily have similar structures (de Amorim et al., 2013).

1.3.4 Periplasmic Haem Binding Proteins

A core component of nutrient uptake in Gram-negative bacteria is the periplasmic binding protein (PBP). These proteins are important in facilitating the dissociation of the haem or iron from high affinity scavenger proteins which require a conformational change to release their ligand. There are three classes of PBP determined by the number

of connecting regions between the two lobes characteristic of this protein family. Class I has three β connector regions, Class II two β, and Class III just one alpha helical connecting region. These connector regions are involved in the conformational changes that occur upon ligand binding.

In *H. influenzae*, the removal of iron from the OM receptor-bound transferrin complex requires the combination of; 1) A possible conformational change to transferrin upon binding to the OM receptor TbpAB, allowing transferrin to release the iron and; 2) the PBP, in this case *H. influenzae's* ferric binding protein A (FbpA), to capture this iron due to its higher affinity for iron than transferrin (Khan et al., 2007). The PBP then acts as a shuttle system to take the iron across the periplasmic space to the ABC transporter at the cytoplasmic inner membrane.

The ferric enterobactin iron uptake system has been described and the siderophore not only binds to the OM receptor FepA with a reported affinity in the nanomolar range, but also to the PBP FepB with higher affinity- a K_d of 30 nM as opposed to 60 nM with FepA, which may be vital in the siderophore crossing the cell membrane (Sprencel et al., 2000). FepB, like ShuT and PhuT, is a class III PBP. A study by Chu *et al* investigated the conformational changes occurring upon ligand binding. Class I and II PBPs display a higher degree of movement in a 'Venus fly trap' like motion upon ligand binding, however Chu *et al* showed this movement was minimal in class III PBPs, and binding was centred around a movable loop region that placed the PBP in an advantageous position for binding to the ABC transporter FepCD (Chu et al., 2014).

The class III haem binding PBP found in *Shigella dysenteriae,* ShuT, had its haem binding mechanism characterised by Eakanunkul *et al*, where it was seen to interact with hemin via a proximal Fe^{3+}-O bond to Tyr-94 (Eakanunkul et al., 2005). The amino acid sequence of ShuT has high levels of identity to predicted haem binding PBPs from other pasteurellaceae species (see Figure 1.6).

Other PBPs that have been identified in *pasteurellaceae* include SapA, a haem binding protein found in *H. influenzae* shown by Mason *et al* to be vital in haem uptake across the cytoplasmic membrane via the transporter protein complex SapBC (Mason et al., 2011). HbpA is another haem binding lipoprotein predicted to reside in the periplasm of

H. influenzae, again involved in the uptake and transport of haem from hemoproteins such as hemopexin and haemoglobin. Its importance is highlighted in a study performed by Morton *et al* where a HbpA mutant was not able to grow as well as the wild type counterparts on haem only medium (Morton et al., 2009). The study also concluded that *H. influenzae* had additional haem PBP's, as the mutant was able to grow on haem based medium albeit at a reduced rate.

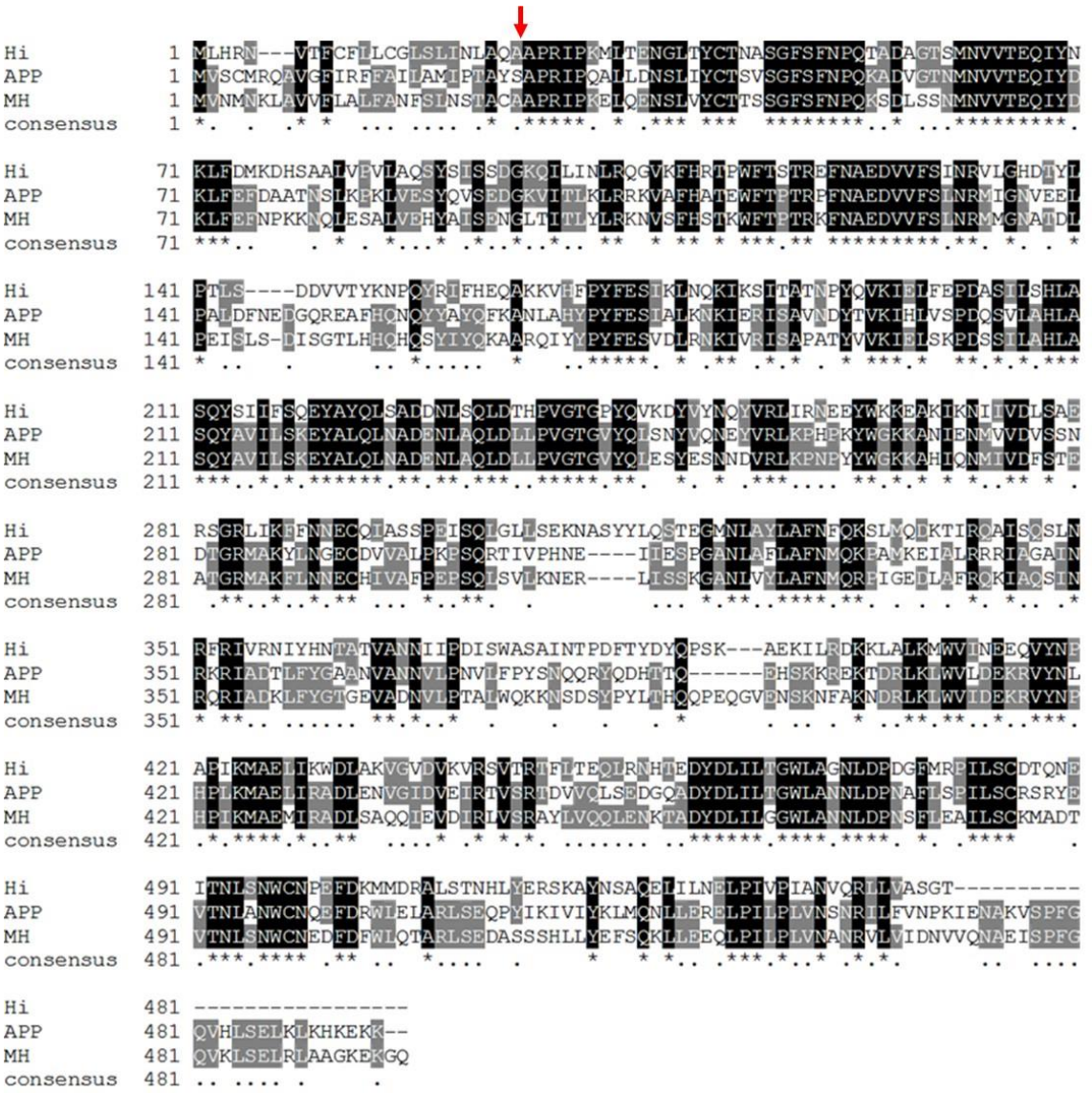


Figure 1.6 Periplasmic Binding Proteins in *Pasteurellaceae*.

The alignment shows the *H. infuenzae* SapA (HI) AAQ12665.1 amino acid sequence with potential homologues identified in *A. pleuropneumoniae* (APP) EFL80116.1 and *M. haemolytica* (MH) WP_042803671.1. The levels of identity between these individual amino acid sequences are between 46.72% and 60.82%. * denotes identical residues in all sequences whereas . denotes identical residues in the majority of sequences. Signal peptide cleavage site indicated by red arrow.

1.3.5 ABC Transporters

ABC transporters are proteins situated on the cytoplasmic membrane of the Gram-negative bacterial cell. They are essential for the active transport of nutrients such as iron or haem into the cytoplasm. The PBP delivers the substrate to the ABC transporter

which in turn provides a channel for said substrate to pass through the inner membrane into the cytoplasm. ABC transporters are generally composed of four subunits; two transmembrane domains and two ATP binding cytoplasmic domains (see Figure 1.7). In the case of *Yersinia pestis*, the PBP HmuT recognises HmuU1 and 2 of the ABC transporter on the cytoplasmic membrane and transfers the haem to a channel between the monomers. ATP drives the PBP shuttling across the periplasm and enables the conformational change, opening the channel between helix 5 in each transmembrane domain (TMD) of the ABC transporter (Woo et al., 2012). The substrate can then cross into the cytoplasm.

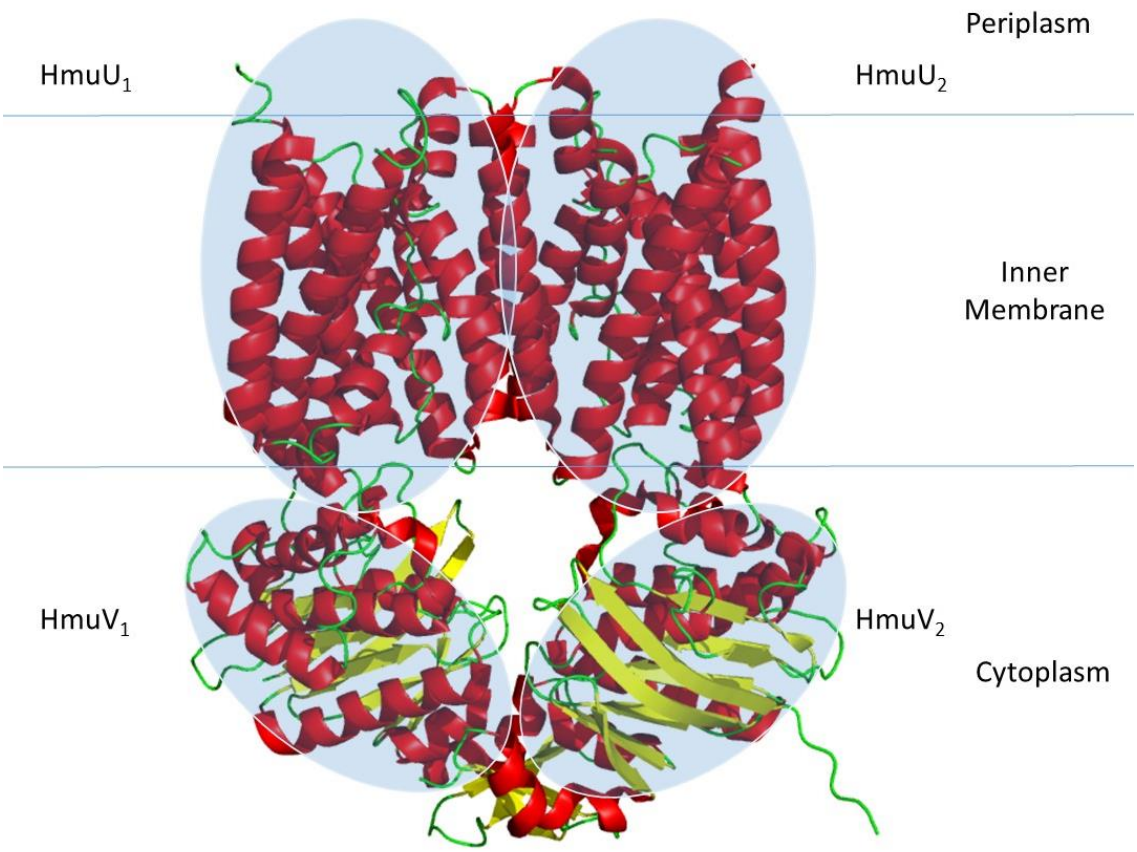


Figure 1.7 ABC Transporter HmuUV in *Yersinia pestis*.

This haem ABC transporter complex structure (PDB code 4G1U) displays the characteristic two transmembrane domain- two ATP binding protein domain structure. A channel opens between $HmuU_1$ and $HmuU_2$ allowing haem to pass through the cytoplasmic membrane. The ABC transporter domains are on the cytoplasmic side of the inner membrane and provide the energy needed to import the haem. Structure taken from RCSB-PBD and rendered in Pymol.

The Sap system in *H. influenzae* is believed to behave in a similar manner to that of *Y. pestis*. The PBP SapA shuttles the haem across the periplasm, delivering it to the transmembrane proteins SapBC. This undergoes a conformational change using the energy derived from hydrolysis of ATP on SapDF, the final components of the ABC transporter complex (Morton et al., 2009). With the BtuCD transporter, the hydrolysis of ATP allows the TMD region to open, accepting the ligand and causing the subsequent release of ADP, which allows the TMD proteins to flip inwards, enabling the ligand to then pass into the cytoplasm (Hollenstein et al., 2007). Both of these uptake systems

rely on type II transporters. While not identical in mechanism of uptake, the Sap and Hmu systems are related to the Vitamin B-12 transporter BtuCD, highlighting the diverse functions of the family of proteins.

1.3.6 Iron and Haem Homeostasis

Whilst nutrient uptake, and in particular iron acquisition, is vital for cellular survival, iron and haem can also be highly toxic to the bacterial cell. It is for this reason that both the iron and haem levels within the cell and the production of proteins involved in the nutrient uptake systems are tightly controlled. Many bacteria such as *N. meningitidis* are capable of synthesising their own haem in a process using multiple enzyme catalysts (see Figure 1.8). This requires exogenous iron to make haem from protoporphyrin IX (PPIX). Not all bacteria have the enzymes required for haem biosynthesis, for example *H. influenzae* lacks the first 6 enzymes in the pathway (Figure 1.8). While the cell can acquire iron for its metabolic needs, it relies on exogenous sources of haem for use in cytochromes and other hemoproteins necessary for cellular functions. White *et al* showed that *H. influenzae* could survive for a short while on haem stores when grown in a haem-limited medium, but growth slowed and only adding exogenous haem back to the enabled cellular replication to resume (White and Granick, 1963).

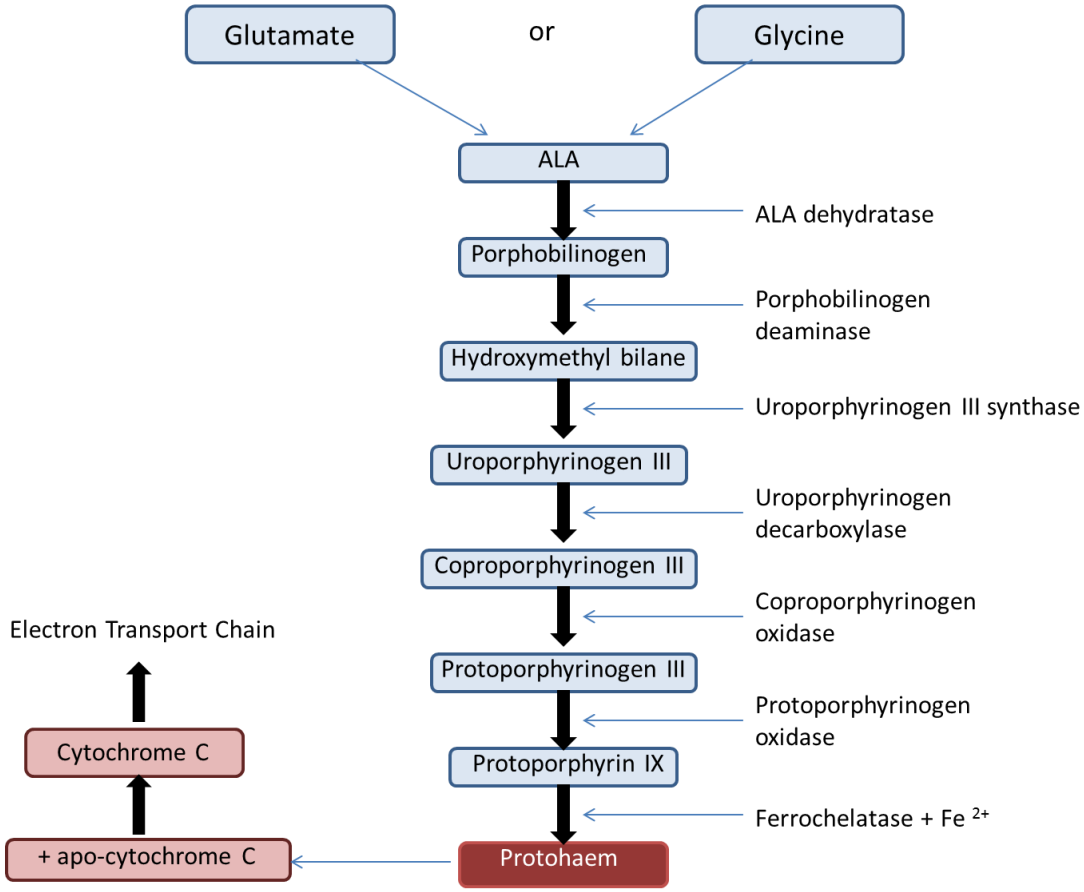


Figure 1.8 Haem Biosynthesis in *N. meningitidis*.

An example of bacterial haem biosynthesis with the completed haem being incorporated into cytochrome C. This pathway is well conserved in many bacteria, animals and plants. ALA - Aminolevulinic acid. Adapted with permission from (Anzaldi and Skaar, 2010).

To avoid cellular toxicity, haem and iron is either utilised in metabolic functions as it is internalised or, if there is an excess, it is sequestered in storage forms within the cell. Iron can be stored in bacterial proteins such as ferritins to avoid the build-up of toxic ROS thereby reducing the potential for DNA and protein damage. When haem is the iron source, the enzyme haem oxygenase (HO) liberates the iron atom by breaking down the protoporphyrin ring leaving CO and biliverdin as by products. A study by Letoffe *et al* identified the proteins YfeX and EfeB as haem-iron liberating agents in *E. coli* (Letoffe et al., 2009). These proteins were subsequently found to be conserved in a wide range of

bacteria, and function to remove the iron from haem by a novel mechanism 'deferrochelation'.

Bacterial haem and iron sensing mechanisms are crucial to adaptation and survival to allow the cell to regulate protein production depending on available resources. The Fec system regulates iron uptake in *E. coli*. The Fec signalling system starts with ferric citrate binding to the OM protein FecA. The periplasmic N- terminal of this interacts with the C-terminal of FecR, an inner membrane protein. FecR then interacts with FecI, which is an extracytoplasmic sigma factor, so called as it is responding indirectly to extracellular stimuli. FecI then acts on the genes associated with the FecABCDE uptake system causing transcription to occur (Enz et al., 2003).

Bacterial iron levels are tightly regulated by variations in gene expression dependant on the cellular environment. Whitby *et al* demonstrated the importance of iron and haem regulation in *H. influenzae* with microarray results showing that approximately 10% of the entire genome is dedicated to acquisition and control of homeostasis of this particular nutrient (Whitby et al., 2006). Iron restriction has been shown to promote the expression of iron acquisition genes such as the ExbBD genes in *N. gonorrhoeae* (Jackson et al., 2010), haemoglobin receptors in APP and general iron transport proteins and permeases in *M. haemolytica* (Klitgaard et al., 2010*)*. Down-regulated genes in *M. haemolytica* were those involved in metalloprotein formation (Roehrig et al., 2007).

The regulation of iron-related gene transcription is widely controlled by the Ferric Uptake Regulator (Fur). The best known function of the protein Fur is as a co-repressor with metals such as iron, cobalt or manganese. The protein dimerises with the help of the metal ion, causing a conformational change. This enables binding of the dimer to a region overlapping the promotor of the target gene known as a 'Fur Box'. The Fur dimer prevents RNA polymerase binding to the promotor and thereby inhibits transcription. When intracellular levels of iron fall, the Fur dimer dissociates and leaves the promoter, and transcription can then occur (see Figure 1.9) (Stojiljkovic and Hantke, 1995). Fur has also been reported to positively regulate transcription when in its monomer form, although the mechanisms are not completely understood. A hypothesis for the mechanism of positive regulation suggests that dimer binding may prevent the

association of other repressors to the DNA, allowing transcription of certain genes (Delany et al., 2004, Yu and Genco, 2012).

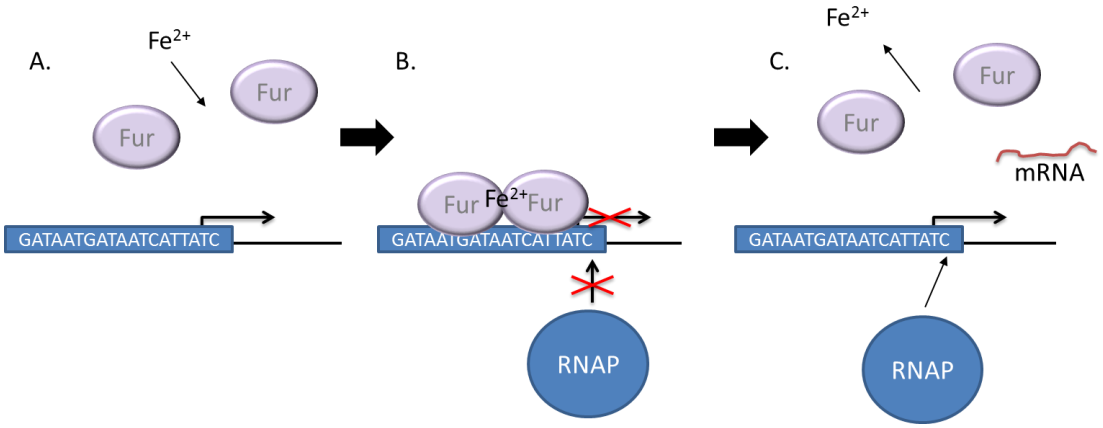


Figure 1.9 Repression of Transcription by the Fur Dimer.

A. In a high iron environment, iron enables dimerization between Fur monomers **B.** The dimer binds the Fur Box (sequence shown Fur Box of *E. coli*) preventing transcription **C.** Cellular iron levels fall, the iron dissociates and transcription can resume. RNAP denotes RNA polymerase.

In specific environmental niches, nutrients may be limited. It is therefore important for the bacterial cells to conserve energy whilst maintaining the ability to acquire sufficient resources to survive. In concentrating the resources available to the vital functions for cellular survival and the appropriate nutrient acquisition, the bacterial cells are giving themselves the best chance of survival in a harsh environment. Fur and other systems of transcription regulation, such as Zur and ArsR, are crucial in the control of these functions, which are also known as bacterial virulence factors.

1.4 Prophylaxis and Treatment

BRD, porcine pleuropneumonia and ovine mastitis all contribute to the disease burden on feedlot animals in the agricultural industry worldwide. BRD causes huge financial problems to the agricultural industry due to the costs of the loss of cattle to BRD, the weight loss during illness due to loss of appetite and the time taken to recover. Morbidity due to infection has been estimated at 20.6% (Faber et al., 1999). Economic costs of this disease are caused by a mortality rate estimated at 6%, decrease in value of infected animals and treatment costs. The decrease in animal value includes a reported $57 lower net profit due to less weight gain in infected animals and lower quality meat (Faber et al., 1999). Treatment costs again have been estimated at $12 per head of cattle in one study performed in the U.S.A, with 81% of this being the costs of the drugs administered. It is telling of the impact of BRD on feedlots that 87% of all treatments administered to cows are aimed at BRD complex (Faber et al., 1999). One of the major etiological causes is *M. haemolytica*, and this pathogen is also a commonly isolated bacteria from sheep suffering from mastitis, which can impact on nursing lambs of infected sheep, causing a further burden to the industry.

Vaccination of cattle prior to introduction into feedlots had proved to be of inconsistent efficacy (Larson and Step, 2012). As a result of this, farmers have turned to antimicrobials, namely antibiotics, as the primary measure to prevent serious losses to their cattle stocks. Due to liberal non-strain specific use of these antimicrobial agents, the rise of antimicrobial resistance is becoming a real problem.

1.4.1 Prevention

Metaphylaxis is used to protect high risk cattle upon entry to the feedlot prior to the development of immunity resulting from vaccination. In the USA, over 73% of cattle received metaphylactic antibiotics upon entry to feedlots in 2011 ((USDA), 2013). Common antibiotics registered for this use include tilmicosin, florfenicol, tulathromycin and ceftiofur. These have to be controlled and prescribed by experienced veterinarians to reduce the chances of the development of antibiotic resistance. Metaphylaxis has been estimated to reduce disease incidence by up to 50% and overall mortality by up to

25% although meta-analysis highlighted that results for these investigations are inconsistent due to large differences in variables (Van Donkersgoed, 1992). While this is impressive, it has been widely suggested that improvements in vaccination protocols could render this unnecessary, thereby decreasing the chance of selecting resistant bacteria.

It can take between 1-3 weeks for calves to develop an immune response post-vaccination, and further booster doses may be required depending upon the vaccine used. There are a wide range of vaccines available against the common viral and bacterial causes of BRD, which can be used in a variety of combinations. Preconditioning of the cattle has been shown to aid in the decrease in disease incidence, however this is not always possible (Taylor et al., 2010).

Cattle vaccinated with killed vaccines had higher morbidity rates than those administered modified live vaccines and 52% of calves vaccinated in one study relapsed after vaccination (Faber et al., 1999). A meta-study conducted by Larson *et al* on over 20 years of data found that vaccines against *M. haemolytica* did not significantly decrease mortality in North American feedlots, although an overall decrease was observed in many of the trials studied (Larson and Step, 2012). The same was found for *H. somni*, although the number of trials assessed was much lower. Overall the trials assessing the efficacy of vaccines used to reduce BRD caused by bacteria were inconsistent, and had many limitations (Larson and Step, 2012). Current vaccines used against the Gram-negative bacterial causes of BRD include killed vaccines, such as Pulmoguard (Prolabs), and combined vaccines such as the newly released in 2014 Titanium 5 + PH-M (Elanco) which covers both viral and bacterial causes. Independent reviews of the latter have not been reported, yet reports from the company show that it does reduce the mortality resulting from *M. haemolytica* infections when compared to an unvaccinated control group (Milliken, 2013).

Phase variation is one recognised mechanism of bacterial survival, whereby changes in expressed proteins can aid host defence avoidance. An example of this is the expression of variations of the PorA protein on the outer membranes of *N. meningitidis*, enabling the bacteria to avoid antibody detection (Tauseef et al., 2013). Phase variation can also

present a problem when considering drug targets; a high level of phase variation in a target protein for example would decrease the efficacy of the drug designed to work via that particular protein. A protein that is necessary for some aspect of cellular survival and proliferation is more likely to be expressed, and can therefore present a more attractive target.

1.4.2 Treatments

Effective treatment of BRD requires knowledge of the specific causes to tailor the therapy appropriately, and to minimise the selection of resistant bacteria. Classes of antibiotics used include macrolides (e.g. Tilmicosin), chloramphenicols (e.g. Nuflor), tetracyclines, cephalosporins (e.g. Naxcel) and fluoroquinolones (e.g. Baytril). The mechanisms of action are to disrupt or entirely inhibit cell wall synthesis in the case of cephalopsporins, which inhibit the peptidoglycan crosslinking by penicillin binding proteins. Protein synthesis is targeted by macrolides and tetracyclins, which prevent the formation of peptide bonds via inhibiting the 50S subunit interactions and the binding of aminoacyl tRNA to the ribosomal subunits, respectively. Fluoroquinolones inhibit nucleic acid synthesis by targeting DNA topoisomerases. Despite efforts to reduce the rise of resistant bacterial strains, reports over recent years have indicated that the number of antibiotics that still provide protection are dwindling.

1.4.3 Antibiotic Resistance

Antibiotics currently used to both prevent and treat bacterial infections involved in BRD have been widely used in agriculture for many years. While attempts have been made more recently to control the indiscriminate usage, there are still instances of inappropriate and over use. Resistance can occur by favourable mutations to targets such as the topoisomerases, preventing antibiotic binding. Limiting the availability of the antibiotic by decreasing membrane permeability or an increase in efflux pumps can also lead to failure. Other mechanisms include the inactivation of antibiotics, such as chloramphenicol, by acetyl transferases. The gene for β-lactamase also confers resistance to penicillin and others in this class of antibiotics by targeting the β-lactam ring leading to inactivation.

Klima *et al* found that of the *Mannheimia* isolates from diseased animals investigated, 37% (Klima et al., 2014a) and 72% (Klima et al., 2014b) in two different studies had some variety of antibiotic resistance, with up to 18% of these showing resistance to tetracycline (Klima et al., 2014a, Klima et al., 2014b). All of the isolates displaying antibiotic resistance in both studies were serotype A1, with the exception of one serotype: A2. In a further study, they investigated the importance of integrative conjugative elements (ICE) in conferring resistance upon bacteria, finding that 88% of isolates from the samples from the USA contained multidrug resistant ICE (Klima et al., 2014b). ICE regions are known mobile genetic elements able to transfer between competent bacteria, which presents a further survival mechanism. A dangerous by-product of this was highlighted in the ability of these regions to transfer between *P. multocida* and *E. coli* (Klima et al., 2014b). This could potentially lead to the spread of resistance not just intra-species but also in bacteria colonising similar ecological niches. Plasmid mediated resistance has also been observed in *M. haemolytica*, with genes conferring resistance to both florfenicol and chloramphenicol found in a particular isolate (Katsuda et al., 2012). Resistance is not limited to *M. haemolytica*, with reports of disease isolates of *A. pleuropneumoniae* harbouring resistance genes to tetracycline. Lower levels of resistance to florfenicol, tilmicosin and ampicillin were also noted in this species (Kucerova et al., 2011).

1.4.4 New Targets

As illustrated by the high cost of treatments for BRD, new targets that can affect numerous causative agents of this multifactorial disease complex could significantly lower this economic burden. Additional significant costs include the prophylaxis and treatment of herd of cattle. To target bacteria in a manner to reduce the likelihood of resistance arising, key components in cell survival are an obvious choice.

A promising vaccine trial used the *A. pleuropneumoniae* Apx toxins against them, utilising a fragment of ApxII as part of a vaccine. Nasal immunization was shown to induce protective immunity against infection by multiple serotypes of *A. pleuropneumoniae* in a murine model (Seo et al., 2013). Using this toxin as a target may function by both neutralising the damaging effect of the toxin alone and preventing the

bacteria from colonising the respiratory and reproductive mucosa by maintaining low numbers of bacteria (Chiang et al., 2009, Seo et al., 2013).

As previously discussed, OM proteins can be involved in adhesion and are also known to be immunogenic. The addition of one of these, PlpE, to vaccines for *M. haemolytica* was found to enhance vaccine efficacy and broaden the specificity to A6 as well as A1 (Confer et al., 2006). There is a need for greater cross-serotype protection amongst *M. haemolytica* vaccines, as the majority target A1. There is a reported increase in the prevalence of other serotypes causing infection, therefore a conserved protein throughout serotypes would be an ideal target.

1.5 The Gly1 Protein Family

1.5.1 Neisserial Gly1 Proteins

The protein gonolysin1 (Gly1) was originally discovered by Arvidson *et al* in *Neisseria gonorrhoeae* while screening the genome for putative bacterial toxins (Arvidson et al., 1999). The *gly1* gene was found by screening clones created in an *E. coli* library containing DNA of *N. gonorrhoeae* and investigating the coding DNA in colonies that displayed some haemolytic activity. Strains of *N. gonorrhoeae* exhibited increased toxicity to human fallopian tube cultures when Gly1ORF1 protein was mutated, and found to reduce the toxicity when wt Gly1ORF1 was added back into the growth medium, although this finding has not been confirmed with subsequent experiments. Gly1ORF1 was discovered at a locus along with a second open reading frame (ORF), with no obvious promoter sequence in between the two genes. The proteins, Gly1ORF1 and Gly1ORF2 were therefore thought to be co-transcribed (Figure 1.10). Gly1ORF1 was the protein of interest in this study. This protein had no known function, and screening in the NCBI database initially revealed no matches, with more recent searches finding only 'hypothetical' proteins with similar sequences. Arvidson *et al* found that this protein was mainly located in outer membrane blebs of the Neisseria along with some associated with the OM itself. The Gly1 protein was not detected in the soluble lysate proteins of *N. gonorrhoeae*, yet was observed in small quantities in outer membrane blebs- 10^{10} cells were required to detect the protein via immunoblot. Although the Arvidson group

hypothesised it to be related to inhibition of cellular toxicity in the fallopian tubes, further work has not confirmed this. Studies by the Arvidson group also showed that the Gly1 proteins were not essential factors in neisserial adhesion or invasion into host tissues *in vitro* (Arvidson et al., 1999).

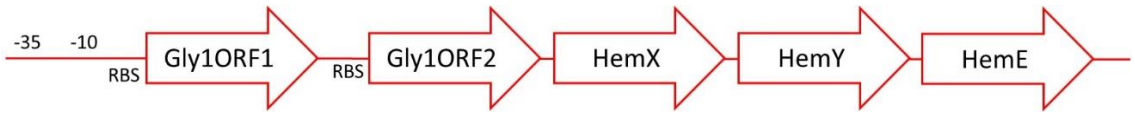


Figure 1.10 Gly1 Locus in *Neisserial* Species

The *gly1ORF1* and *gly1ORF2* genes are located adjacent to each other in the genome. No promoter binding region was discovered between the two genes however a predicted ribosomal binding site (RBS) was discovered. Gly1ORF2 had some similarity to *E. coli HemD*. Other haem biosynthesis genes were found downstream of the Gly1 locus.

The second protein of the locus, Gly1ORF2 was also entered into the NCBI BLAST database search, and the results here showed that it shared a degree of similarity to *E. coli's HemD*, a protein involved in haem biosynthesis (see Figure 1.8). More recent searches of the NCBI database have further suggested that this is the neisserial HemD, with updated annotations of the genomes showing a very high degree of conservation of this gene and its locus between neisserial species and strains. Other flanking genes of Gly1 also appeared to be involved in haem biosynthesis (see Figure 1.10 and 3.5 in Chapter 3) which lead to the hypothesis that Gly1 is involved some way in haem uptake or metabolism.

The *Neisseria gonorrhoeae* Gly1ORF1 protein sequence was entered into the NCBI BLAST database to search for potential family members in other bacterial species. The results of this search are ever increasing as more and more genomes are added to the database. The Gly1 protein was found to be highly conserved throughout neisserial species including *Neisseria meningitidis* and *Neisseria lactamica* to name a few. These bacteria are found in human hosts, often as commensals in the respiratory tract in the case of *N. lactamica* and *N. meningitidis*, however disease can occur when *N. meningitidis* moves to the lower respiratory tract or enters the bloodstream.

Sathyamurthy conducted experiments with the wild type and a ΔGly1 mutant strain of *N. meningitidis* strain MC58, the mutant created with the insertion of an erythromycin cassette within the *gly1* gene. These experiments showed that while the wild type bacteria were able to grow in medium containing either iron, haem or haemoglobin, the ΔGly1 knock out strain was unable to grow in medium containing haem or haemoglobin alone. When purified Gly1 protein was reintroduced to the medium, growth was restored in a dose dependent manner. The ΔGly1 mutant was able to grow on medium containing inorganic iron (Sathyamurthy, 2011).

Further work with the *N. meningitidis* Gly1 protein also showed interactions with the neisserial HmbR (Weirzbicka, 2014) suggesting that Gly1 has a central role in iron uptake from haem sources. Interactions with erythrocytes were also demonstrated, with Gly1 causing deformation of the red blood cells (Sathyamurthy, 2011) thought to be via transient protein interactions (Weirzbicka, 2014). Characterisation of this interaction was suggested to be via the erythrocyte membrane protein EBP 4.2, initially shown by Yeast-2-hybrid experiments (Meadows, 2005). Sathyamurthy *et al* demonstrated interactions using fluorescently labelled neisserial Gly1 incubated with a mouse cell line expressing the membrane protein (Sathyamurthy, 2011). These interaction results could not be replicated using purified EBP 4.2 in Gly1 binding experiments performed using the BLItz (ForteBio), which was suspected to be a result of incorrect folding or protein instability (Weirzbicka, 2014).

In addition to this potential role in nutrient uptake, serum bactericidal assays were performed which showed bactericidal activity of serum antibodies raised against the *N. meningitidis* Gly1 protein (Sathyamurthy, 2011, Weirzbicka, 2014). These data supported the idea that anti-Gly1 antibodies could be used as a potential vaccine candidate, although further experiments would be required in an animal model of neisserial infection to verify this.

Studies of antibacterial agents focussing on iron acquisition in the form of haem uptake include targeting HasA of *P. aeruginosa*. This investigation showed that HasA was able to bind synthetic metals rather than haem, and in doing so on a haem only medium *P. aeruginosa* was unable to grow (Shirataki et al., 2014). Not all vaccine candidates have

shown cross-serotype protection, for example using HgbA of *H. ducreyi* along with an adjuvant only conferred protection against those isolates expressing one class of receptor (Fusco et al., 2010). Highly conserved target proteins throughout serotypes would enable a cross-serotype vaccine, and potentially even cross-species.

1.5.2 Potential Homologues- The Gly1 Protein Family

Potential homologues were discovered in *Actinobacillus pleuropneumoniae*, *Histophilus somni* (formerly *Haemophilus somnus*) and *Mannheimia haemolytica* (formerly *Pasteurella haemolytica),* among others. Of particular interest is the recent addition of a putative homologue in *P. multocida*, another bacteria commonly isolated in BRD along with *M. haemolytica* and *H. somni*. Of these species, *M. haemolytica* boasts 12 different serotypes, *P. multocida* 5 and *A. pleuropneumoniae* 15. *H. somni* is not classed into serotypes that are widely used, and just two strains have their complete genome sequences available as of August 2015.

Further NCBI BLAST searches revealed potential protein homologues of the Gly1 protein within a number of Gram-negative bacteria, ranging from highly similar neisserial species such as *N. lactamica* as previously mentioned to *S. enterica IIIa* and *E. tarda* - reptile and fish pathogens, respectively.

1.5.3 Project Aims at the Outset

Due to the economic importance of the diseases caused by these animal pathogens, the members of this possible new 'gonolysin1- like' protein family may be of interest as they are present in all available sequenced neisserial genomes at time of writing. This, coupled with prior research showing neisserial Gly1 proteins conveying complement-mediated bactericidal activity, could present Gly1 proteins as potential novel targets for vaccines or therapy.

The initial aims of the project were to investigate potential Gly1 family proteins found within a number of animal pathogens. The proteins in question were of unknown function and structure, with only the ORFs highlighted within the genomes in the NCBI database as an indication that they existed. Therefore the proteins from *M. haemolytica*,

A. pleuropneumoniae and *S. enterica arizonae (IIIa)* were chosen to undergo characterisation experiments to discover whether their structures and/ or functions were similar to that of the neisserial Gly1 proteins previously investigated. The initial hypothesis was that the Gly1 protein family is involved in haem- iron acquisition in gram-negative bacteria based on the neisserial Gly1 studies conducted in the Sayers group.

Specifically, we set out to:

- Identify putative sequence homologues in different bacterial species
- Discover the prevalence of Gly1 homologues in *M. haemolytica* isolates
- Screen these isolates to analyse sequence conservation within *Mannheimia* species
- Express and purify Gly1 homologues and investigate the structural similarities of these proteins
- Determine haem binding ability of these proteins of interest.

2. Materials and Methods

2.1 Cloning

2.1.1 Preparation of Chemically Competent Cells

Axenic cultures of *E. coli* M72, BL21 (DE3), BL21 (DE3) RIPL or XLI Blue cells (all lab strains, see Appendix for genotypes) were grown on LB agar plates overnight at 30°C (see section 2.2.4). One colony was used to inoculate 10 ml of LB medium, which was incubated overnight at 30°C. The 10 ml culture was added to 100 ml LB and incubated at 30°C until an A_{600} of 0.4 was reached. The cell suspension was centrifuged at 335 x *g* for 10 minutes at 4°C. The cell pellet was re-suspended in 25 ml 100 mM $MgCl_2$. The suspension was again centrifuged at 335 x *g* for 10 minutes, 4°C, and the cell pellet gently re-suspended in 5 ml of 100 mM ice cold $CaCl_2$. A further 45 ml of 100 mM ice cold $CaCl_2$ was added and this solution was incubated on ice for 1 hour. The cells were again centrifuged at 335 x *g* for 10 minutes at 4°C. This pellet was re-suspended in 6 ml 100 mM ice cold $CaCl_2$ plus 15% glycerol and immediately frozen in aliquots at -80°C.

2.1.2 Transformation of Chemically Competent Cells

Competent cells that had been grown in 0.1% glucose enriched LB medium were defrosted on ice and 50 µl added to pre-chilled micro-centrifuge tubes. The control tubes contained 1. competent cells only, and 2. an empty vector relevant to the transformation. To each subsequent aliquot, 2.5 µl of diluted plasmid was added. Samples were incubated on ice for 1 hour followed by a 1 minute heat shock at 42°C. The cells were immediately returned to ice for 5 minutes. A 50 µl aliquot of each sample was plated onto an LB agar plate containing ampicillin and 0.1% glucose, and incubated at 30°C overnight.

2.1.3 Mini Prep of Plasmid DNA

Cells containing the plasmid (XLIB for pET-21-a and pUC19 storage followed by BL21 (DE3)) for subsequent protein overexpression, and M72 for pJONEX4 and pJONEX4 C-his) were grown in LB medium plus ampicillin overnight at 30°C. For XLIB and BL21 (DE3)

cells, LB was enriched with 0.1% glucose. Of this, 6 ml was centrifuged, the supernatant was discarded and the cell pellet used for a DNA mini-prep (Omega Bio-Tek USA, Plasmid Mini-Prep Kit). For midi-preps (Qiagen Ltd, UK), 100 ml overnight cultures were grown, of which 50 ml was used.

2.1.4 Double Digest of Gene and Plasmid

Restriction sites compatible with the multiple cloning site in pJONEX4 (EcoRI and HindIII), for pJONEX4 C-his (EcoRI and BamHI), for pUC19 (EcoRI and BamHI) and for pET21-a (NdeI and HindIII) were incorporated into the gene of interest using custom designed primers (see table 2.1) and were designed to allow over expression from either the lambda, lac or T7 promoter depending on the vector of choice (see Table 2.1). For the double digest, an initial single digest was carried out using approximately 40 ng of vector from the mini-prep with 10 units of each restriction enzyme plus 2 µl 10 x buffer and dH_2O to a final volume of 20 µl. An undigested control was included containing only 10 ng vector plus dH_2O to a final volume of 5 µl. Samples were incubated for 1 hour at 37°C. Of each single digest, 15 µl was transferred into a single tube to give a final volume of 30 µl. A further 10 units of each restriction enzyme was added to this and all samples were incubated for 1 hour at 37°C. A 1% agarose gel containing ethidium bromide (final concentration of 1 µg/ml) was prepared and 5 µl of each sample, mixed with 1 µl 6 x loading dye, and either a 100 bp or a 1 kb DNA ladder (both NEB), was run at 4 V cm^{-1} for approximately 45 minutes. A large well containing 25 µl (approximately 250 ng) of the double digest mixture was also included for gel extraction. The wells containing approximately 50 ng DNA were UV imaged and marks made to enable extraction of the larger band of DNA without UV exposure. Gel Extraction was performed using a Qiagen Gel Extraction kit. The APP Gly1 gene was synthesised by Eurofins MWG Operon in the pEX-A-new plasmid to enable restriction digest with EcoRI and HindIII, allowing for direct cloning into the pJONEX4 vector.

Target gene	Forward sequence	Reverse sequence
MHA_1205	5'-TTTCGAATTCTAGAGGAAACAAAAATGCGTAAATTATTAGTAATT-3' (EcoRI)	5'-TGTTGGATCCCTAAAGTATTCATCAAATGAACATC-3' (BamHI)
MHA_0579	5'-AAAAGAATTCAAAGGAAACTAACATATGAAAAAAATGGC-3' (EcoRI)	5'-TTAAGGATCCCGGCTGCAAATCCAGATTTTCTCATCA-3' (BamHI)
MHpara	5'-AAAGGAAAACATATGAAAAAATTGATGATATTCGC-3' (NdeI)	5'-TCTTAAGCTTTTACTATTTTTGAGGTGTGCAACCTTT-3' (HindIII)
PhuT	5'-AATACATATGAGAACCTTACTTTTTTCTTT-3' (NdeI)	5'-TATAAAGCTTATTTTAAACAAGGTTGAATAGT-3' (HindIII)
HmbR	5'-TAAAGGATCCAGAGGACTAAAATATGCAAATAAAGCAAA-3' (BamHI)	5'-AAAAGAATTCATTTACAAAAAATGGGCCAAAATTA-3' (EcoRI)

Table 2. 1. Primers Used for Cloning. Restriction enzymes in brackets and recognition sites underlined, start codon highlighted in yellow.

2.1.5 Ligation of Gene into Vector

Ligation of the Gly1 insert into the plasmid was performed using an approximate 1:3 molar ratio of vector: insert. Samples were incubated in ligation buffer and 5 units of T4 DNA Ligase (both Fermentas) and made up to 10 µl with dH_2O. The reaction was left on ice overnight. A 1% agarose gel was run to check ligation. The complete plasmid was then transformed back into the appropriate competent cell type (see section 2.1.2). After transformation, 50 µl of cells were plated (see section 2.2.4). These included a negative control of cells containing no vector, a control for self-ligation with vector and ligase alone, and insert and ligase alone. The experimental cells contained different ratios of the ligation of vector and insert and a positive control of the undigested vector. The plates were grown at 30°C. Colonies were grown up in 6 ml LB cultures (see section 2.2.5) and the DNA extracted using the Omega plasmid mini-prep kit as before. The DNA was sent for sequencing using either M13 primers applicable for both of the pJONEX4 vectors or T7 primers for the pUC19 and pET-21a vectors to assess the quality of the

cloning. Sequencing was performed by the University of Sheffield Core Genomics Facility, and sequences were analysed using FinchTV 1.4.0 (Geospiza, Inc.; Seattle, WA, USA; http://www.geospiza.com).

The *M. haemolytica* Gly1 paralogue was cloned from a 50 ng stock of genomic DNA (a gift from Intervet Innovation GmbH, Schwabenheim, Germany). The periplasmic haem binding protein PhuT and the haemoglobin receptor (HmbR) were also cloned from this genomic DNA.

2.2 Protein Overexpression and Purification

2.2.1 Cell Culture- Buffers and Reagents

Lennox Broth (LB) medium consisted of 1% tryptone w/v, 0.5% yeast extract (Oxoid) w/v, 0.5% NaCl (Acros Organics) w/v plus 1.5% Bacto™ Agar (BD) w/v for agar plates. The 4YT (yeast- tryptone) medium consisted of 3.2% tryptone, 2% yeast extract and 0.5% NaCl. The ampicillin and carbenicillin (Melford) were both diluted to a final concentration of 100 µg/ml in cell culture media. Cell lines used were *E. coli* K- 12 M72, *E. coli* BL21 DE3 and *E. coli* XLI Blue (all lab strains). Cell wash buffer consisted of 25 mM Tris·HCl pH 8, 2 mM EDTA pH 8 and 100 mM NaCl.

2.2.2 Heat Induced Expression System of pJONEX4

M. haemolytica Gly1 *MHA_1205* and *MHA_0579* were cloned into the plasmid pJONEX4-C-His (Sayers and Eckstein, 1991). To enable specific selection, the plasmid contained a gene coding for beta lactamase, which confers ampicillin resistance. A λP_L promoter controlled transcription of Gly1 via a heat inducible system. The plasmid was transformed into M72 cells to enable over expression via a heat inducible cI857 repressor on a defective prophage.

2.2.3 IPTG Induced Expression System of pET21-a(+)

An alternate system for protein over expression used IPTG induction of the vector pET-21a(+) (Novagen). Also containing the beta lactamase gene conferring ampicillin resistance, cell lines containing this vector (BL21 DE3 for over-expression and XLI Blue for storage at -80°C) were grown in the same medium as pJONEX4 containing M72 cells,

but at a lower temperature of 25°C to avoid problems with solubility. Induction was instigated by the addition of between 0.5 mM and 1 mM IPTG to a cell culture of A_{600} 0.4. The BL21 DE3 cells were grown to this density in the presence of 0.1% glucose to prevent any leaky expression of T7 RNAP from the lac promoter.

2.2.4 LB Agar Plates

LB medium was made (see section 2.2.1) and 1.5% w/v agar added. The solution was autoclaved, cooled, and had appropriate antibiotics added as required prior to pouring. For XL1 Blue cell culture glucose was also added to a final concentration of 0.1% v/v prior to pouring. To plate the bacteria, a colony-sized volume of stock cells containing the plasmid (stored at -80°C, 20% glycerol) was streaked on an LB agar plate using a sterile loop and flame. The plate was incubated at 30°C overnight for pJONEX4 containing cells and 37°C for pET-21a(+), pUC19 and pEX-a-new vectors to get individual colonies. From liquid cell cultures, 50 µl was plated, spread and left to dry prior to inverting and incubating at a temperature appropriate for the plasmid.

2.2.5 Small to Medium Scale Cell Culture

Ampicillin was added to 3- 10 ml LB medium to final concentration of 100 µg/ml. A single colony was used to inoculate the medium and left overnight at 30°C or 37°C, shaking at 200 rpm. From the overnight culture, 1 ml was added to 100 ml LB with ampicillin (final concentration of 100 µg/ml) and allowed to either grow to the appropriate A_{600} for small scale expression, or to grow overnight for use as a feeder for large scale culture.

2.2.6 Large Scale Cell Culture

A 5 l fermenter was used to provide improved aeration of the cells in the hope of achieving a higher yield of protein. To 5 l of 4YT in the fermenter, 400 ml of overnight culture was added along with carbenicillin (see section 2.2.1) and 0.75 ml antifoam (Sigma). The fermenter was aerated (two volumes per minute) and stirred at 300 rpm constantly at a temperature of 25°C until it reached an A_{600} of 2. A 1 ml un-induced sample was taken, centrifuged and the supernatant and pellet separated then frozen.

For cells containing pJONEX4 vectors, the temperature was then increased to 42°C for a 3 hour induction. After induction, the temperature was reduced to 26°C, and fermentation continued for 13 hours. For cells containing pET21-a(+) and pUC19, IPTG was added as described in section 2.2.3, with the temperature remaining at 25 °C. Post induction, a 1 ml sample was again taken, centrifuged and supernatant and pellet separated then frozen. The remaining culture was centrifuged at 9000 x *g,* 20 minutes and the supernatant stored at 4°C. The cell pellet was re-suspended in 200 ml cell wash buffer. The suspension was centrifuged at 15000 x *g* for 30 minutes at 4°C, supernatant discarded and pellet re-suspended in 200 ml of cell wash buffer minus EDTA. This was centrifuged and the supernatant discarded, and pellets frozen at -80°C.

2.3 Protein analysis via SDS-PAGE

2.3.1 SDS-PAGE- Buffers and Reagents

A 13% SDS-PAGE gel was used throughout to enhance resolution of lower MW bands e.g. Gly1. The resolving gel contained 13% (w/v) acrylamide, 25 mM Tris-Bicine pH 8, 0.1% SDS, 0.5 mg/ml APS and 0.016% (v/v) TEMED. The stacking gel contained 4% (w/v) acrylamide, 100 mM Tris·HCl pH 6.8, 0.08% SDS, 0.4 mg/ml APS and 0.2% TEMED. Sample loading dye consisted of 0.2% SDS, 40% glycerol, 75 mM Tris pH 8, a few crystals of bromophenol blue (Sigma), 5 mM EDTA (BDH) pH 8, with fresh dithiothreitol (DTT) (Melford) crystals added prior to use. The cell lysis buffer consisted of 25 mM Tris·HCl pH 8 and 1 mM EDTA pH 8.

2.3.2 Sample Preparation SDS-PAGE Gel

The cell pellet from the 1 ml samples was re-suspended in lysis buffer normalised using the A_{600} value at which induction was initiated i.e. (A_{600} x 100)/2, unless the culture was small-scale and no A_{600} reading was taken, in which case 100 µl of lysis buffer was uniformly used. To the cell suspension, SDS solution was added to a final concentration of 1%; the sample was mixed by vortex and then boiled for 3 minutes at 100°C. A 1:1 ratio of loading dye (see section 2.3.1) was added to the sample and the solution was boiled again for 3 minutes at 100°C. A 10 µl sample was loaded on an SDS-PAGE gel. Cell

supernatants were mixed 1:1 with loading dye, boiled once at 100°C for 3 minutes and loaded onto the SDS-PAGE gel.

2.3.3 Acetone Protein Precipitation

Protein in the supernatant samples was precipitated out using the Thermo Scientific Acetone Precipitation Protocol. A 1:4 ratio of sample: acetone (pre-chilled to -20°C), was mixed using a vortex and incubated for 1 hour at -20°C. The sample was micro-centrifuged at 13,000 x *g* for 10 minutes, the acetone discarded, and the pellet left to air-dry for 35 minutes. The pellet was re-suspended in 10 µl deionised water (dH_20) and 10 µl loading dye. The sample was then heated for 3 minutes at 100°C and run on an SDS-PAGE gel.

2.4 Protein Purification- Buffers and Reagents

2.4.1 Nickel Chelate Column Buffers

Nickel column wash buffer consisted of 25 mM Tris·HCl, 500 mM NaCl and 5% glycerol v/v pH 8. This buffer was also used for dialysis of samples prior to running on the nickel column. Nickel column loading buffer consisted of 25 mM Tris·HCl pH 8, 500 mM NaCl, 5% glycerol v/v and 20 mM imidazole. Nickel column cleaning buffer consisted of 3 M guanidine-HCl (GuHCl) (Acros Organics), 5 mM EDTA and 50 mM Tris·HCl pH 8.

2.4.2 Ion Exchange Buffers

For Q columns, low salt Q buffer consisted of 25 mM Tris·HCl, 1 mM EDTA, 5% glycerol v/v and 50 mM NaCl pH 8. Samples were dialysed into low salt Q buffer prior to running through the Q column. High salt Q buffer consisted of 25 mM Tris·HCl, 1 mM EDTA, 5% glycerol v/v and 1 M NaCl pH 8. Buffers for the SP column consisted of 20 mM KH_2PO_4, 20 mM K_2HPO_4, 5% glycerol, 2 mM EDTA and 50 mM or 1 M NaCl for the low and high salt elutions, respectively.

2.5 Protein Purification

2.5.1 Bacterial Cell Pellet Lysis

The frozen cell pellet was re-suspended in lysis buffer in a volume of 5 ml per gram of pellet at room temperature. Cell lysis buffer consisted of 200 mM NaCl and 25 mM Tris·HCl pH 8. Once fully re-suspended, lysozyme (Sigma) was added to achieve a final concentration of 200 µg/ml. PMSF (Sigma) was added to a final concentration of 23 µg/ml. As the cell suspension became viscous, sodium deoxycholate (Acros Organic) was added to a final concentration of 1 mg/ml. The cell suspension was then sonicated (Jencons Vibra Cell) in 30 second bursts at 20% amplitude until thinned, left on ice for 20 minutes, and sonicated again in 30 second bursts to ensure full disruption of the cell membranes. The lysed cell suspension was centrifuged at 20000 x *g* for 20 minutes at 4°C, and the supernatant transferred to a clean container. To this, ammonium sulphate (Sigma) was added to achieve a final concentration of 3 M and left on a roller for minimum 1 hour at 4°C for protein precipitation. The suspension was centrifuged at 40000 x *g* for 20 minutes at 4°C and the supernatant discarded. The pellet was re-suspended in dialysis buffer and prepared for dialysis.

2.5.2 PhuT Solubilisation

The *M. Haemolytica* periplasmic haem binding protein, PhuT, was over expressed using the pET21a vector after transformation into BL21 (DE3) *E. coli* cells. A small quantity of the over expressed protein leaked out into the culture medium and was concentrated down for purification. The protein in the pellet fraction however, was found to be insoluble after cell lysis. To resolubilise the protein, KP buffer was added to re suspend the post-sonication pellet, and GuHCl was added in powder form to a final concentration of 7 M. The sample was left mixing for a minimum of 1 hour at 4°C. DTT to 1 mM was added to prevent any oxidative damage to the cysteines. The sample was centrifuged at 40000 x *g* for 20 minutes and the supernatant was then separated from the non-soluble pellet fraction. The supernatant was then slowly dialysed into Q buffer pH 8 in preparation for ion exchange chromatography. This allowed re-solubilisation of approximately half of the insoluble protein.

2.5.3 Nickel Chelate Column Chromatography

To purify the polyhistidine-tagged Gly1 proteins from the cell-free extract a 10 ml column containing 50% HIS-Buster Nickel Affinity Gel (Amocol) slurry was used. The column was equilibrated with 25 ml sodium acetate followed by 25 ml of wash buffer. A 50 µl sample of the cell-free extract was retained and stored at 4°C. The remaining cell free extract was diluted 1:2 with loading buffer and loaded onto the column using gravity flow. The flow through was collected in fractions, and the column washed with 25 ml load buffer. The protein was then eluted using 10 ml washes of 50 mM, 100 mM and 200 mM imidazole and collected in two 5 ml fractions. A final elution with 20 ml 500 mM imidazole was collected in two 10 ml fractions. The column was washed once more with wash buffer, cleaned with cleaning buffer (see section 2.4.1), and each flow through collected. All washes were stored at 4°C. The column was stored in 20% ethanol. To recharge the nickel beads, 100 mM $NiCl_2$ was loaded onto the column twice. The samples were analysed on an SDS-PAGE gel.

2.5.4 Ion Exchange Chromatography

To further purify the protein of interest, anion and/or cation exchange chromatography was performed with either a Q column or an SP column respectively (GE Healthcare). The 1 ml or 5 ml Q column was equilibrated by washing with 10 column volumes (c.v.) low salt Q buffer, 10 c.v. high salt Q buffer, and 10 c.v. low salt Q buffer again. Following buffer exchange dialysis (see section 2.6), the eluted protein from the nickel column was loaded onto the Q column at 0.5 ml/min and the flow through collected. A 50 µl sample of the load was retained and stored at 4°C. The column was washed with 10 ml low salt Q buffer at a rate of 0.5 ml/min. The proteins were then eluted using a 20 ml NaCl gradient ranging from 50 mM to 1 M, followed by 5 ml of high salt Q buffer and fractions of 1 ml were collected. The column was then washed with 10 ml 3 M GuHCl followed by 10 ml dH_2O and stored in 20% ethanol at 4°C. The above method was performed for SP columns using the low and high salt SP buffers referred to in section 2.4.2.

2.6 Protein Analysis

2.6.1 Dialysis

Dialysis tubing (Spectra/Por 1 Dialysis Tubing, 6-8 kDa MWCO, Spectrumlabs) was boiled in 2 mM EDTA and 100 mM $NaHCO_3$, followed by 100 mM $NaHCO_3$ alone, then washed thoroughly with dH_20. The solution to be dialysed was added to the tubing and secured. The dialysis bags were incubated in a minimum of a 40 fold dilution of the required buffer at 4°C for 4 hours, the buffer changed, and the dialysis continued overnight in a further minimum of a 40 fold dilution to allow equilibrium to be reached.

2.6.2 Protein Quantification

To measure protein concentration, a Bradford Assay was performed. The Protein Assay Reagent (Biorad) was diluted 2:7 in dH_2O. Standards of 10, 8, 6, 4 and 2 µg of BSA (Sigma) in a 100 µl volume plus 900 µl diluted Protein Assay Reagent were measured at A_{595} in triplicate. The protein sample was diluted into a final volume of 100 µl in dH_2O to fall within the limits of the standards. The A_{595} was measured in triplicate and the protein concentration deduced graphically from comparison with the known standards. Alternatively, the Pierce BCA assay kit was used following the user guidelines.

2.6.3 Concentration of Protein Samples

To increase the concentration of protein, the solution was centrifuged at 657 x *g* for 20 minutes at 4°C in a 5000 MWCO Viva Spin tube (Sartorius Stedim Biotech) until the required volume was achieved. Alternatively, acetone precipitation was used (Ras, 2009) for concentration of protein found in the sample supernatant. Briefly, a 1:4 dilution of the protein sample in ice cold acetone was mixed by vortex and stored at -20°C for 1 hour. The sample was then centrifuged at 13,000 x *g* for 10 minutes, the supernatant removed and the pellet left to air dry at room temperature. For SDS-PAGE analysis, the pellet was then resuspended in 10 µl dH_2O before adding 10 µl SDS-PAGE loading dye, boiling for 3 minutes and running on a SDS-PAGE gel.

2.6.4 Circular Dichroism Spectroscopy

The circular dichroism (CD) spectra were collected using the Jasco J-10 Spectropolarimeter. The settings used are outlined in Table 2. 2. Proteins were measured between approximately 0.09 mg/ml and 0.2 mg/ml. All proteins were diluted in 1 x PBS buffer, with a baseline taken of the same 1 x PBS buffer. The ellipticity was calculated by subtracting the baseline from the protein values, with the spectra analysed using the following equation used to calculate molar residue ellipticity:

$$[\theta]\ MRW = \frac{\theta \times 100 \times Mr}{C \times D \times N_A}$$

Condition	**Setting**
Sensitivity	100mdeg
Start	300 nm
End	190 nm
Data pitch	1 nm
Scanning mode	Continuous
Scanning speed	20 nm/min
Response	8 sec
Band width	1 nm
Slit width	1 nm
Accumulations	5

Table 2. 2 Circular Dichroism Settings

2.6.5 Polyhistidine Tag Cleavage

Enterokinase (polyhistidine tagged) (Biorbyt Ltd, UK) was diluted to 0.1 and 0.04 units per microliter in storage buffer (200 mM NaCl, 20 mM Tris·HCl, 2mM $CaCl_2$). Target proteins were dialysed into 1 x PBS prior to use and 20 µg was used per 50 µl reaction. Each reaction was composed of 5 µl 10 x reaction buffer (20 mM $CaCl_2$, 500 mM $NaCl_2$, 200 mM Tris·HCl pH8), 20 µg target protein, 5 µl diluted enterokinase and dH_2O to make up the 50 µl volume. Reactions were left for 24 hours at 25°C, then analysed on SDS-PAGE gels for cleavage. Once cleavage was confirmed, a small scale nickel pull-down was performed to remove both the cleaved polyhistidine tag and the enterokinase. Briefly, 50 µl of nickel beads were washed and centrifuged as with the nickel column (see section 2.5.3). The reaction was added to the beads, incubated at room temperature for one hour then centrifuged at 5000 x g in a desk top centrifuge for 5 minutes. The supernatant was carefully removed and a sample analysed on an SDS-PAGE gel alongside the protein before enterokinase cleavage.

2.6.6 Antibody Production and Purification

Purified APP protein resuspended in 1 x PBS to a concentration of 200 µM was injected into 14 week old BALB/C mice, 100 µl per injection at d0, d14 and d28 over a period of 8 weeks. Tail bleeds were performed at 3 weeks and 6 weeks and a terminal bleed was taken at 8 weeks after confirming the presence of the antibody in the serum.

MHA_0579 C- His and MHpara antibodies were created by Generon using 1 mg/ml pure protein in a 60 day immunisation and bleed protocol to ensure a response was elicited. Antibodies were subsequently purified using affinity chromatography columns specific for each antibody (ThermoFisher AminoLink Coupling resin). Briefly, corresponding proteins were immobilised onto equilibrated columns, washed with Binding Buffer and the antisera added. The columns were then washed again with Binding Buffer to remove non-specific antibodies and other proteins and the protein specific antibodies were then eluted using Elution buffer. Antibody presence was confirmed using western blotting.

2.7 Bioinformatic Analysis

The database used in this project for discovery of genomic information and protein sequences was the NCBI BLAST site (Altschul et al., 1997)

http://blast.ncbi.nlm.nih.gov/Blast.cgi. To search for potential homologues of the initial neisserial proteins, their amino acid sequences were used to screen genomes uploaded in the database using an initial BLASTp search and a maximum of seven iterations. Potential protein homologues were chosen from these searches that had an 'FSC', 'FMC' or 'YSC' region near the N terminal, a C terminal cysteine and a predicted N terminal signal peptide sequence of between 19 and 24 amino acids in length. Signal peptide prediction was performed using SignalP (Choo et al., 2009) http://www.cbs.dtu.dk/services/SignalP/. The SABLE server (Adamczak et al., 2005) http://sable.cchmc.org/ was used to predict the secondary structure and Phyre2 (Kelley et al., 2015) to predict unsolved tertiary structures of the *M. haemolytica* Gly1 homologues. PsortB V3.0 was used to predict the cellular location of proteins (Yu et al., 2010).

2.8 X-ray Crystallography

Proteins were purified as mentioned in section 2.4, concentrated to a minimum of 10 mg/ml, and buffer exchanged into 10 mM TE buffer in the case of the *M. haemolytica* Gly1 paralogue, and MHA_0579 C-His was crystallised in PACT D6, which is 0.1 M MMT (DL-malic acid: MES: Tris base) buffer pH 9 and 25% (w/v) PEG 1500 respectively. The X-ray crystallography on the *M. haemolytica* proteins was performed by Jason Wilson, Dept. of Molecular Biology & Biotechnology, The University of Sheffield. Screening and optimisation was carried out on each protein. Additional screening of the MHA_0579 C-His protein was attempted in the presence of hemin. The hemin soak was performed by adding small amounts of solid hemin to the drop with an acupuncture needle**.** Crystallography on SARI C-His was initially also attempted by Rick Salmon also of MBB at TUOS.

2.9 Hemin Interactions

2.9.1 Hemin Agarose Binding Assay

Hemin immobilised on agarose beads was used to assess if bacterial proteins of interest interacted with hemin. Stock concentrations of proteins were diluted to approximately 2 µg/ 20 µl in 1 x PBS plus 20 µg/ 20 µl BSA. The bovine hemin agarose beads (Sigma

Aldrich) were washed 3 times with 0.5 ml 1 x PBS 500 mM NaCl, centrifuged and the pellet re-suspended in 1 x PBS to achieve 50% bead slurry. To 20 µl aliquots of hemin beads, 10 µl of each protein stock solution or a further negative control of 10 µl 1 x PBS BSA only was added. The samples were mixed on a vortex and left rolling at room temperature for 1 hour. The samples were centrifuged at 5000 x *g* for 2 minutes and the supernatant and hemin bead pellet separated. The beads were washed three times with 0.5 ml 1 x PBS 500 mM NaCl and centrifuged. All samples were re-suspended in 10 µl loading dye, boiled for 3 minutes at 100°C and loaded onto an SDS-PAGE gel (2.3). Densiometry analysis was performed using ImageJ (Rasband, 1997-2012) software, and the Students *t*-test was performed to generate statistical data. If proteins interacted with the hemin they would be found in the 'pellet' fraction associated with the hemin beads.

Proteins used included the Gly1 proteins of interest, recombinant human growth hormone (polyhistidine-tagged, a kind gift from Dr. Pradhananga, TUOS), human flap endonuclease (FEN) and PhuT proteins. *N. meningitidis* Gly1 was used as a positive control, as haem binding had previously been shown with this protein in the Sayers lab (Sathyamurthy, 2011). Lysozyme was used as a negative control.

2.9.2 Haemoglobin Agarose Binding Assay

Haemoglobin (immobilised on agarose beads) was employed using the same method as the hemin agarose bead binding assay in section 2.9.1. In addition to this, glutaraldehyde was used to make any binding permanent between the Gly1 and haemoglobin to capture the binding in the event of a reversible reaction.

2.9.3 Hemin Absorbance Assay

Hemin (Sigma) was dissolved in 100 mM NaOH, and diluted with 1 mM Tris·HCl pH 8 to produce 0.1 mM hemin stock solution at pH 7.5. The hemin was diluted to a working concentration of 25 µM, with Gly1 protein concentrations varying. Samples consisted of 250 µl Gly1 or other protein plus 250 µl hemin, along with a hemin-PBS control, and protein only. The samples were made up to 1 ml with 500 µl 1 x PBS. The absorbance

spectra were scanned from 260 nm- 500 nm wavelength, using PBS only as the baseline. Each spectrum is representative of a minimum of three repeats.

2.9.4 Hemin Peroxidase Activity

A 96 well microtitre plate with a maxisorb surface was coated with 0, 2, 4, 6, and 8 µg of protein made up to 100 µl with 1 x PBS buffer, and incubated at 37 °C for 2 hours. The plate was then washed with 200 µl 1 x PBS three times. 20 µg of hemin/ 100 µl 1 x PBS was added to each of the experimental wells, with 1 x PBS only or hemin to wells with no initial protein coating added to the control wells. The plate was incubated for 1 hour at 37 °C, and washed a further three times with 200 µl 1 x PBS. TMB (Sigma) was added at 50 µl per well and the reaction stopped after 45 minutes to allow sufficient colour change to occur. The reaction was stopped with 50 µl 1 M HCl and absorbance was measured on a plate reader at 450 nm.

2.10 PCR Screening

Isolates 234, 235, 292, 296, 311, 351, 362, 411, 414 and 428 were kindly donated by Dr Miranda Kirchner at the AHVLA in Surrey, England. These were received on Dorset's Egg Slope medium and re-streaked onto Columbia blood agar (CBA) plates made with defibrinated horse blood. In addition, a further 19 bacterial isolates were generously gifted on CBA plates by Professor Andrew Rycroft at the Royal Veterinary College, Potters Bar (see Table 2.3). The plates were incubated overnight at 37 °C and 5% CO_2. Single colonies were resuspended in 20 µl TE buffer and boiled at 95 °C for 5 minutes. The samples were then centrifuged for 1 minute at 13,000 x *g* on a desktop centrifuge and the supernatant transferred to a fresh Eppendorf tube. This supernatant was stored at -20 °C until required for PCR reactions. Isolates were stored in cryo-protective tubes at -80 °C.

Strain	Date (If known)
M. haemolytica **A10** (T10)	5/75
M. haemolytica **A12**	5/75
M. haemolytica **p28**	22/08/80
M. haemolytica **A1**	5/75
M. haemolytica **A7**	5/75
M. haemolytica **A9**	5/75
M. haemolytica **CO2392**	
M. haemolytica **CO0600**	
M. haemolytica **CO4178**	
M. haemolytica **CO3983**	
M. haemolytica **CO3916**	
M. haemolytica **Gilly**	2007/2008
M. haemolytica **CO0353**	
M. haemolytica **MH4 A8**	5/75
M. haemolytica **MH8 A5**	5/75
M. haemolytica **MH9 A2**	5/75
M. haemolytica **A6**	76/1
M. haemolytica **CO2008**	
M. haemolytica **CO1250**	
234	Ovine, 05/2012
235	Bovine, 05/2012
292	Ovine, 09/2012
296	Ovine, 09/2012
311	Ovine, 09/2012
351	Ovine, 05/2012
362	Bovine, 10/2012
411	Ovine, 11/2012
414	Ovine, 11/2012
428	Ovine, 12/2012

Table 2.3 Isolates Investigated.

The first 19 received from the RVC courtesy of Prof Rycroft, the latter 10 from the AHVLA courtesy of Dr Kirchner. The above isolate names have been shortened to that in bold for the purpose of this thesis. Serotypes were available for 9 of the 19 isolates acquired.

For the PCR, Q5 Hot Start Polymerase (NEB) was used throughout. Initial screening reactions were carried out in 25 µl reaction volumes, and subsequent reactions for DNA sequencing were in 50 µl volumes. Typical 25 µl reactions consisted of 5 µl NEB Q5 reaction buffer, 200 µM dNTP's, 10 pmol forward primer, 10 pmol reverse primer, approximately 10 ng DNA and 1 unit of Q5 Hot Start Polymerase, made up to 25 µl with

dH_2O. Reaction conditions were as follows: Initial denaturation at 95 °C for 5 minutes (1 cycle), followed by 25 cycles of denaturation at 95°C for 1 minute, annealing at variable temperatures for 1 minute, elongation at 72°C for 1 minute. The final elongation was carried out at 72 °C for 2 minutes. Samples were stored at 4 °C until separated on an agarose gel.

The variable annealing temperature was optimised to each particular primer set as stated in the primer summary table. The PCR reaction was then analysed using a 1% agarose gel. Any positive reactions were PCR purified using the Qiagen PCR purification kit and sent to the core facility for sequencing. Sequences were again analysed using FinchTV 1.4.0 (Geospiza, Inc.; Seattle, WA, USA; http://www.geospiza.com). Sequences were aligned with the expected gene sequence based on the primers used with the Boxshade server. Any sequences that were showing poor identity, or were partial gene sequences for isolate identification purposes were screened using the NCBI BLAST database to discover the identity of the sequence and or isolate. ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) was used to analyse the impact of any substitutions or deletions throughout the sequences.

Name	Target gene	Forward	Reverse	Tm °C
Set A	MHA_0579	5’-CAACGCACCAAAAGGAAACT-3’	5’-ATGGTGGGTTAATGCTTTGC-3’	53
Set D	MHA_1205 or COI_1584	5’-GGCCCTAGTGCCATACAGG-3’	5’-CTGTCAGAAATCGCCAGGTT-3’	54
Set J		5’-CTGTCAGAAATCGCCAGGTT-3’	5’-GATTTACCTACCGAGCATTGT-3’	54
Set L		5’-TTTACCTACCGATCTTTGT-3’	5’-CAAACTGGTGATATGAATGCTTT-3’	50
Set M		5’-CGCCAAGTTGATAATAATG-3’	5’-CGAGCTTTGTATTTTTTAAC-3’	48
Set H	MHpara	5’-ACTCACTTTCTAAACAATAGC-3’	5’-CGCTTGTTTCTGGATAGTA-3’	50
PhuT	PhuT	5’-GTTAACACTATATCATTTTAATTGAA-3’	5’-AAATGGGTACAAATCACTAAG-3’	50
16S		5’-GCTAACTCCGTGCCAGCAG-3’	5’-CGTGGACTACCAGGGTATCTAATC-3’	59
23S		5’-GACGGAAAGACCCCGTGAACCT-3’	5’-GGCAAGTTTCGTGCTTAGAT-3’	53

Table 2. 4. Primers Used for Screening.
Primers used in cloning HmbR were also incorporated into the genomic screening process. Annealing temperature used is stated (Tm).

3 Structural Properties of Gly1 homologues

3.1 Introduction

The neisserial Gly1 proteins comprise a novel tertiary structure dubbed the 'beta-diablo fold', which consists of eight anti-parallel beta strands forming a barrel-like structure. The Gly1 homologues from *N. gonorrhoeae* and *N. meningitidis* have been crystallised showing identical structures. The proteins crystallised as dimers with their larger alpha helical residues facing towards each other, and it was thought that haem binding of the *N. meningitidis* Gly1 occurs in a 2:1 protein to haem ratio (Sathyamurthy, 2011).

A number of species within the *Pasteurellaceae* family of bacteria contain genes that give rise to so called 'hypothetical' proteins of interest- potential Gly1 homologues with varying levels of amino acid sequence identity with the original highly conserved neisserial Gly1 proteins. In order to assess whether these proteins are indeed homologues despite the divergence, sequence analysis and structural studies were performed using a combination of online predictive software and laboratory based methods.

The term 'homologue' in its broadest sense describes genes or protein functions of common ancestry. Paralogues are genes that may be found in the same genome as the original homologue, but differ from the original gene or protein in sequence and/ or function. Protein families can be highly conserved, such as histones or mammalian flap endonucleases, or can share as little as 10- 15% sequence identity, as seen between chicken egg and T4 lysozymes for example, whilst retaining a common function. While a high percentage sequence identity can be a good clue to homology, low sequence identity does not rule it out, and conversely, some proteins with 30- 40% identity may not necessarily be related.

The sequence alignment in Figure 3.1 highlights a collection of 'hypothetical' proteins similar to the neisserial Gly1 protein found in a wide range of bacterial species. The lower the percentage identity between the amino acid sequences, the harder it is to confirm homology without analysis of structure and function.

Candidates for Gly1 homologues found through NCBI BLAST searches have been observed in numerous Gram-negative bacterial genome sequences with varying degrees of amino acid sequence identity (see Figure 3.1). More and more potential homologues are appearing as the online databases grow, with the figure below showing just a fraction of those discovered to date. Due to the vast number of species containing a possible *gly1* gene homologue, for the purposes of this investigation the main focus is on those from *M. haemolytica*, with a parallel investigation into those found in *A. pleuropneumoniae* and *S. enterica arizonae IIIa*.

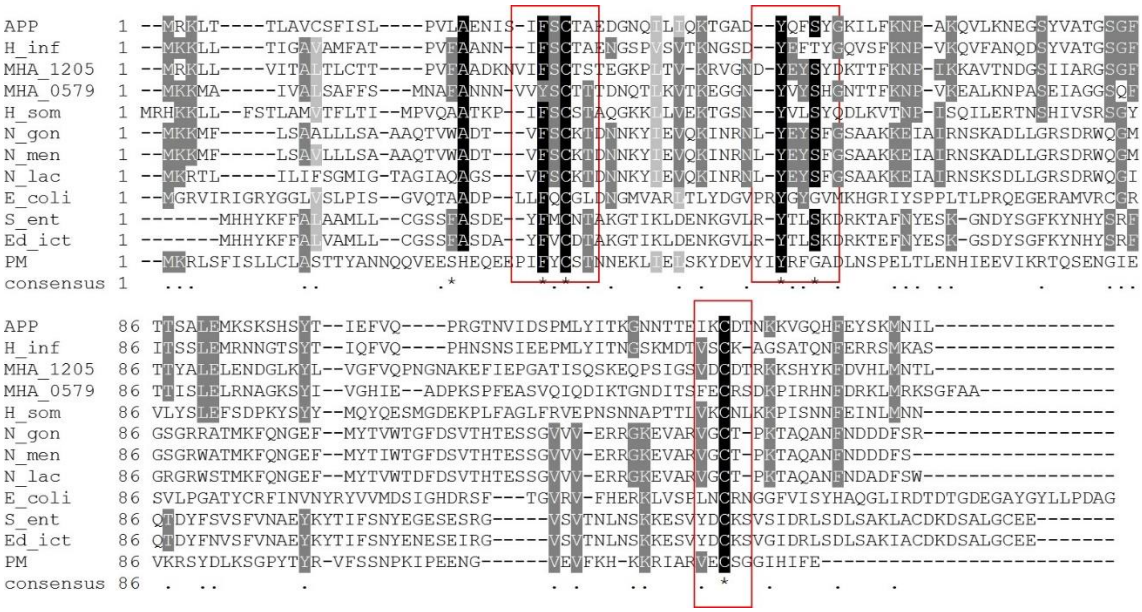


Figure 3.1. Multiple Sequence Alignments of Gly1 Homologues from Different Species.

Alignment depicting similarities between identified Gly1 proteins. Accession numbers are as follows *A. pleuropneumoniae* (APP) WP_005607867.1, *H. influenzae* WP_005653509.1, two *M. haemolytica* proteins, *H. somni* YP_719978.1, *N. gonorrhoeae* ACF29206.1, *N. meningitidis* NP_273818.1, *N. lactamica* WP_003710927.1, *E. coli* EIL56827.1, *S. enterica* YP_001573309.1, *Edwardsiella ictaluri* YP_002932900.1, *P. multocida* (PM) WP_046339897.1 Key regions are the FSC, YEYS (including variants) and the C terminal cysteine (highlighted).

3.2 Results

3.2.1 Homologues and Flanking Genes

In the first instance, identification of genomes available for bacteria of interest containing putative *gly1* genes was imperative. At the time of writing the number of genomic sequences of *M. haemolytica* available on the NCBI database was 19. These differ in their origins; the majority were isolated from bovine hosts, with two from ovine, and one each from porcine and cervine hosts. They are either of unknown serotype or one of the three major worldwide serotypes i.e. those most commonly isolated (see Table).

The potential Gly1 homologues discovered using the initial neisserial Gly1 protein sequence span species from *Actinobacillus pleuropneumoniae,* a pig pathogen, to *Salmonella enterica IIIa,* which lives in reptile and amphibian hosts, to *E. tarda,* a fish pathogen. Higher sequence identity suggests increased likelihood of similar protein function. The neisserial Gly1 proteins share very high sequence similarity, with different neisserial species exhibiting above 98% identity throughout. This level of identity and even the presence of the gene decreases for other species, for example database screening found only one strain of *S. enterica* with a good candidate to be a Gly1 homologue present - *S. enterica arizonae IIIa* which is a less common strain. Additionally *A. pleuropneumoniae* (APP) Gly1 is found only in one strain but 3 assemblies out of 22 available genome assemblies within the database- serovar 6, strain FEMØ (RefSeq I.D GCF_000178555.1). There are 21 genome assemblies of *M. haemolytica* available to screen on NCBI database and hundreds of *S. enterica* genomes, however, just one isolate of *S. enterica arizonae (IIIa)* (RefSeq I.D GCF_000018625.1).

Strain	Serotype	Origin	Pathogenicity
MhBrain2012	A1	Bovine (Brain)	Disease
MhSwine2000	A1	Porcine (Lung)	Pneumonia
M42548	?	Bovine	?
D193	A1	Bovine (lung)	Pneumonia
PHL213	A1	Bovine (lung)	Pneumonia
Serotype A2 str. Bovine	A2	Bovine	Pneumonia
Serotype A2 str. Ovine	A2	Ovine	Pneumonia
D171	A2	Bovine	Pneumonia
D35	A2	Bovine	Pneumonia
Serotype A1/A6 str. PKL10	A1/A6 cross reactive	Deer (Spleen)	Pneumonia
USDA-ARS-USMARC-185	A6	Bovine (nasal swab)	?
Serotype 6 str. H23	A6	Bovine (nasopharynx)	BRD
D38	A6	Bovine (lung)	Pneumonia
D174	A6	Bovine (lung)	Pneumonia
USMARC_2286	? Putative commensal	Bovine (nasopharynx)	Asymptomatic
D153	A1	Bovine	Pneumonia
USDA-ARS-USMARC-183	A1	Bovine	?
USDA-ARS-MARC-184	A2	Bovine	Commensal?
MH10517	?	Ovine (lung)	?

Table 3.1 Genomic Sequences of *M. haemolytica* on NCBI Database.

The genomic sequences listed above were obtained from the NCBI database. They are all full genome sequences, with the assembly level varying from complete genome, to scaffold, to contig. Accession numbers are as follows: MhBrain2012 **GCF_000443225.1,** MhSwine2000 **GCF_000443185.1,** M42548 **GCF_000376645.1,** D193 **GCF_000443205.1,** PHL213 **GCF_000153645.1,** Serotype A2 str. Bovine **GCF_000176275.1,** Serotype A2 str. Ovine **GCF_000176255.1,** D171 **GCF_000427275.1,** D35 **GCF_000443085.1,** Serotype A1/A6 str PKL10 **Genbank accession only available GCA_000584935.1.,** USDA-ARS-USMARC-185 **GCF_000349785.1,** Serotype 6 str H23 **GCF_000341635.1,** D38 **GCF_000443105.1,** D174 **GCF_000422095.1,** USMARC_2286 **GCF_000439735.1,** D153 **GCF_000422145.1,** USDA-ARS-USMARC-183 **GCF_000349765.1,** USDA-ARS-MARC-184 **GCA_000819525.1,** MH10517 **GCF_000931975.1.** All data taken from the NCBI database.

Gly1 homologues in *M. haemolytica* share between 34% and 99% identity with each other, and overall just around 20% identity with the original neisserial Gly1 (see Table 3.3). The *M. haemolytica* Gly1 paralogue (MHpara) shares only 13% identity with the homologues of the same species, however, the MHpara protein (or slight variations of this) has been found in all *Mannheimia* genomic sequences on the NCBI database at the time of writing. APP Gly1 is most similar to that found in the human pathogen *H. influenzae*, but shares around 30% identity with the *M. haemolytica* Gly1 proteins. The *Salmonella* Gly1 protein, SARI, is again most similar to a *M. haemolytica* Gly1, however, the level of identity is lower at between 16- 25%.

Due to the large number of identical proteins found in *M. haemolytica* isolates on the database, the name of one has been used to represent the others present to avoid confusion. The representative protein name is highlighted in bold at the top of each column in Table 3.2. Interestingly, although there are three identical APP Gly1 proteins listed in the PSI-BLAST search, two exist in the same isolate, one in contig4, one in contig29. Genome analysis showed identical flanking regions and these therefore were identified as overlapping contiguous sequences. The third is named as a reference protein for 'Actinobacillus species' which is apparently a multispecies protein. This could mean that only one *A. pleuropneumoniae* strain carries this protein or that the genome has been incorrectly analysed.

APP (hypothetical) WP_005607867.1 *Actinobacillus species*	**MHA_1205 EDN74135.1 *M. haemolytica* PHL213**	**COI_1584 EEY09775.1 *M. haemolytica* serotype A2 str. OVINE**	**mh0603 ABG89179.1 *M. haemolytica* Serotype A1**	**L280_13705 EPZ27028.1 *M. haemolytica* MhBrain2012**	**MHA_0579 EDN73549.1 *M. haemolytica* PHL213**	**SARI (SARI_04389) YP_001573309.1 *S. enterica arizonae* str. RSK2980**	**Putative adhesin (MHpara) AGK03253 M42548 (F)**	Adhesin KIX31695 MH10517 **(F)**
	D650_16360 YP_007666413.1 USDA-ARS-USMARC-183	Hypothetical WP_006251605.1			D650_18560 YP_007666633.1 USDA-ARS-USMARC-183		Adhesin EPZ03517 D38 **(F)**	
	MHH_c22400 YP_007884037.1 M42548	D648_11200 YP_007668632.1 USDA-ARS-USMARC-185			D648_9020 YP_007668414.1 USDA-ARS-USMARC-185		Adhesin AGQ38308 D171 **(F)**	**Adhesin EPY99953.1 *M. haemolytica* D35 (I)**
	J451_13885 YP_008224091.1 D174	L278_09385 EPY99626.1 D35			MHH_c15060 YP_007883313.1 M42548		Adhesin AGQ40883 D174 **(F)**	Adhesin WP_032848124.1 **(I)**
	F382_13645 YP_008235527.1 D153				J451_12635 YP_008223844.1 D153		Adhesin AGQ25325 D153 **(F)**	Adhesin WP_006252140.1 **(I)**
	N220_08230 YP_008339090.1 USMARC_2286				N220_04720 YP_008338402.1 USMARC_2286		Adhesin AGR75796 USMARC_2286 **(F)**	Adhesin AGI31915 USDA-ARS-USMARC-183 **(I)**
	F388_01724 EME04507.1 serotype 6 str. H23				F388_00421 EME04732.1 serotype 6 str. H23		Possible adhesin EDN75149.1 PHL213 **(F)**	Adhesin AGI35981 USDA-ARS-USMARC-185 **(I)**
	L279_08800 EPZ02514.1 D38				L279_12960 EPZ01215.1 D38		F388_12512 EME02394.1 serotype 6 str. H23 **(F)**	Putative adhesin EEY09063.1 serotype A2 str. OVINE **(I)**
	L277_12690 EPZ24426.1 D193				L277_13080 EPZ24311.1 D193		Adhesin EPZ30550.1 D193 **(F)**	Putative adhesin EEY12244.1 serotype A2 str. BOVINE **(I)**
	L281_12465 EPZ26685.1 MhSwine2000				L281_12860 EPZ26326.1 MhSwine2000		Adhesin EPZ26914.1 MhSwine2000**(F)**	
	L280_13960 EPZ26969.1 MhBrain2012				L280_13305 EPZ27491.1 MhBrain2012	**HW40_12515** KIX27265.1 MH10517 *	Adhesin EPZ29268.1 MhBrain2012 **(F)**	**Adhesin AJE08768 USDA-ARS-USMARC-184[A]**

Table 3.2. Protein Labelling. This table shows the **reference protein** used to represent itself and other identical proteins found throughout this thesis (**bold**). Protein name, accession number and strain are listed where known. MHpara is the name used for the so called adhesin throughout this thesis. The second adhesin lane has an I66F polymorphism compared to those in the first column. The final **adhesin** has a frame shift. HW40_12515 has two amino acid substitutions Q118K and D115S when compared to MHA_0579.

	Mhpara	MHA_1205	MHA_0579	MH0603	L280_13705	COI_1584	NM	NG	NL	HS	APP	HI	SARI	PM
MHpara		12.97%	12.97%	13.75%	14.50%	12.97%	16.54%	16.41%	16.03%	12.21%	12.98%	11.57	16.79%	11.38
MHA_1205			35.90%	99.30%	85.60%	99.30%	17.04%	17.52%	14.29%	25.35%	33.58%	39.26	23.91%	15.45
MHA_0579				35.90%	34.00%	36.60%	24.06%	23.13%	16.43%	23.78%	29.93%	34.33	16.90%	15.57
MH0603					85.60%	98.60%	16.79%	19.12%	13.57%	25.35%	33.58%	39.26	23.91%	15.45
L280_13705						86.40%	20.93%	21.54%	14.39%	23.48%	31.06%	36.43	25.00%	14.05
COI_1584							17.04%	15.71%	14.29%	26.06%	33.58%	40.74	23.91%	19.38
NM								97.84%	81.29%	21.21%	21.05%	21.37	17.52%	26.19
NG									80.71%	20.15%	19.40%	21.21	18.12%	26.98
NL										12.86%	13.14%	21.97	15.94%	27.78
HS											27.01%	24.63	20.86%	22.4
APP												49.63	14.60%	13.6
HI													15.04	15.7
SARI														13.14
PM														

Table 3.3. Identities Between Gly1 Homologues.

The above table highlights the similarities in amino acid sequence between the *Mannheimia* Gly1 homologues and those in a selection of other bacterial species. Accession numbers as previously stated for reference proteins. NM = *N. meningitidis* (EFV62744.1), NG = *N. gonorrhoeae* (EFF39185.1), HS = *Histophilus somni* (WP_011609916.1), NL = *Neisseria lactamica* (WP_003710927.1) and HI = *H. influenza* (WP_005653509.1). PM = *Pasteurella multocida*, SARI = *S. enterica arizonae*. Boxes highlighted in red show percentage identity between the two proteins of ≥ 80%. Those highlighted in green show percentage identity of ≥ 30% and ≤ 79.9%. Those in blue share ≥ 20% and ≤ 29.9%. Further homologues in bacterial species exist in addition to those mentioned here.

The criteria used to identify the potentially homologous proteins was based on key featured believed to be involved in the tertiary structure of the neisserial Gly1 proteins. It consisted of 1) an amino acid sequence length of between 130- 145, 2) an N terminal signal peptide of between 19- 24 amino acids in length as predicted by the programme signalP (Petersen et al., 2011), 3) a cysteine close to the C terminal and an FSC region around amino acids 25 to 29, depending on the homologue, and 4) a significant beta strand content. An additional fairly well conserved region noted in the majority of homologues is the YEYS at approximately position 50- 55. This method was fairly inclusive, with a number of potential homologues having either an YSC or FMC region substituting the FSC motif.

Figure 3.2 shows the *N. gonorrhoeae* Gly1 protein structure with the key residues used in identification (highlighted), where they form parts of the antiparallel beta strands and a C terminal turn. The *N. meningitidis* Gly1 dimer structure, when analysed, was found to have a cavity in between the two monomer units. This contains a positive inner region due to the presence of lysine and arginine residues and a negative outer area containing asparagine and glutamate residues. This was identified as a putative haem binding region. Mutations in this region were not found to affect the ability of the Gly1 protein to interact with haem, showing binding must occur at a different site (Sathyamurthy, 2011). The mutated residues did not include those utilised to identify these putative Gly1 homologues. Conserved regions such as those mentioned above are likely to be of functional/ structural importance across homologues.

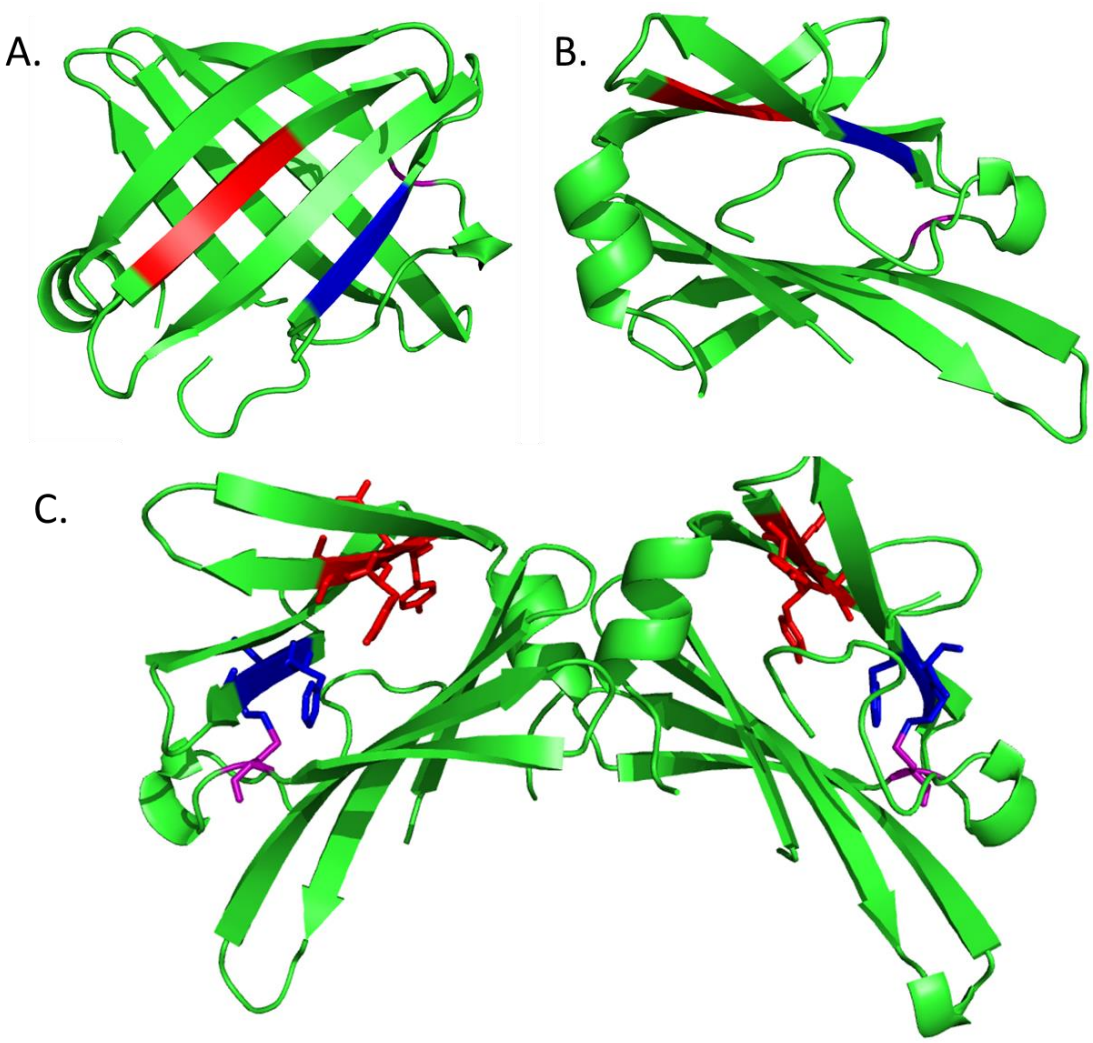


Figure 3.2. *N. gonorrhoeae* Gly1 Structure Showing Key Residues Used in Identification of Homologues (as solved by Arvidson *et al* 2003.)

Panels A and B show the locations of the FSC (Blue), YEYS (Red) and C- terminal cysteine (Purple) regions used in identification of further homologues in two orthogonal views. Panel C shows the amino acid structures as sticks of these regions within the dimer formation that neisserial Gly1 is believed to exist in. Rendered in PyMol from a pdb file supplied by Dennis Arvidson, Michigan State University.

The Gly1 proteins found in *M. haemolytica* are highly conserved, as shown in Figure 3.3, with just MHA_0579 and the MHpara protein standing out due to the lower level of sequence identity. The presence of an N terminal signal peptide was predicted in all homologues and the 'longer' paralogue.

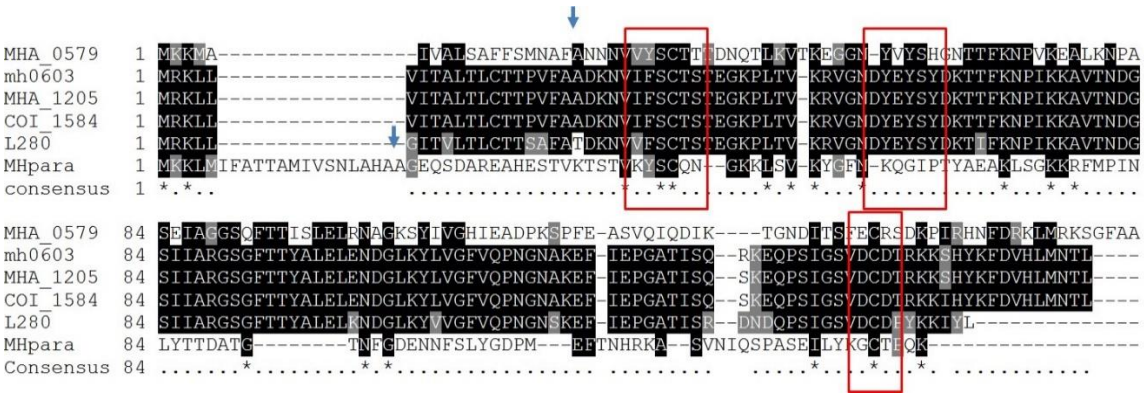


Figure 3.3. *M. haemolytica* Multiple Sequence Alignment.

Alignment showing the regions of similarity between *M. haemolytica* Gly1 homologues discovered to date. Predicted signal peptidase cleavage sites shown with blue arrows. Conserved residues mentioned above highlighted in red boxes.

The *M. haemolytica* paralogue (MHpara) was identified using a PSI-BLAST search of a neisserial Gly1 protein after several iterations. It does not have the same level of sequence similarities or key residues of the initial Gly1 homologues, however, the location in the genome is of interest. Three different paralogues have been found to date (see Figure 3.4) and all are situated in an area of the *M. haemolytica* genome surrounded by genes involved in haem uptake. This is in contrast to the neisserial Gly1 proteins which are adjacent to genes involved in haem biosynthesis. The *Mannheimia, Actinobacillus, Histophilus* and *Salmonella* homologues are flanked by a number of different genes, many of which are of unknown function (Figures 3.5 and 3.6).

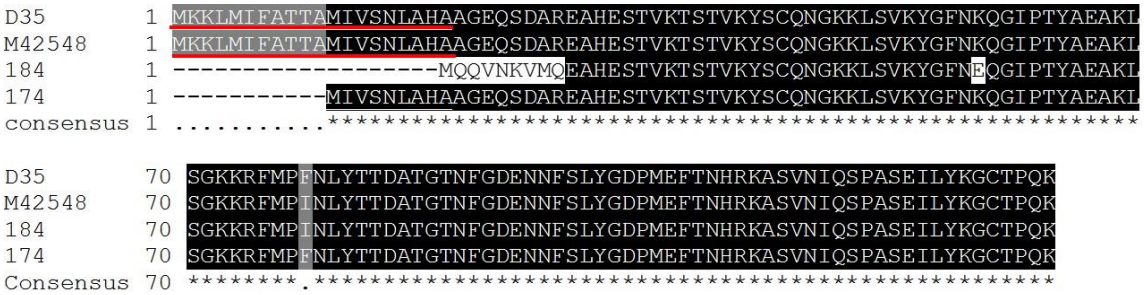


Figure 3.4. Alignments of *M. haemolytica* Gly1 Paralogues.

The predicted signal peptide is underlined in red in the top two sequences. All proteins are labelled as 'adhesin' on the database, therefore an example of the *M. haemolytica* isolate in which they were found has been used to differentiate. Sequence 174 is representative of those suspected to be incorrectly annotated on the database, with identical signal peptides to D35 and M42548 found on closer inspection of the genomes. Sequence 184 has a shorter N terminal region due to the deletion of two DNA bases.

Figure 3.4 shows an alignment of the MHpara proteins. The top two sequences in this figure were annotated on the database as having an additional eleven amino acids at the N terminal position comprising the majority of the signal peptide compared to the lower two. Those starting 'MIVSN' were not predicted to have an N terminal signal peptide when analysed with signalP. Closer inspection, however, revealed the MHpara gene sequence of the shorter protein to be identical to the longer protein upstream of the proposed methionine start codon, suggesting the methionine nominated in the database was incorrect. An area rich in A's and G's around 6 bases upstream of the longer proteins' start codon indicated a possible ribosomal binding site (AAAGGAAA) (Shine and Dalgarno, 1974). This therefore suggests that the different protein lengths are a result of the annotation software used to predict translated proteins within genomes on the database, rather than there being different proteins. The correct sequences were identical to those with the signal peptide in Figure 3.4. The only difference discovered between the two longer proteins is at amino acid position 78, where either a phenylalanine or isoleucine is present.

The third homologue of the MHpara protein, which had an alternate N terminal region that was not predicted to be a signal peptide, was found in isolate 184 and is discussed in more detail in Chapter 6. The PSI-BLAST search also revealed that proteins similar to

the proposed paralogue were found in other species, such as *M. varigena* (WP_025247386.1), *P. bettyae* (WP_005760720.1) and *B. trehalosi* (WP_015431624.1).

The protein PSI-BLAST search identifies proteins similar to the reference protein, in this case *N. meningitidis* Gly1, revealed that the potential Gly1 proteins were all of unknown function, annotated as 'hypothetical proteins'. Previous studies with the *N. gonorrhoeae* Gly1ORF1 protein showed that it was not essential for bacterial survival, and hypothesised that it was involved in regulation of host cellular toxicity caused by the bacteria (Arvidson et al., 1999). Weak haemolytic activity was also observed in wild type cells over-expressing this protein, which was conferred to *E. coli* cells by transformation of Gly1ORF1 in an overexpression system in this investigation. No conclusions were drawn from this study regarding the function of Gly1 other than that further studies need to be conducted. In the absence of published information regarding the possible function of this protein family, we examined the flanking genes of the neisserial Gly1 locus. As described in the operon hypothesis by Monod (Jacob and Monod, 1961), the genomic locus can occasionally provide clues as to the function of the protein transcribed from genes such as the *lac* operon, and more recently in the flagella locus in *A. caviae* which contains five genes involved in flagellar formation (Gryllos et al., 2001).

The flanking genes of the neisserial *gly1* gene in *N. meningitidis* and *N. gonorrhoeae* (Figure 3.5) are identical and code for proteins that are involved in haem biosynthesis, for example the gene supposed to be HemD- *gly1OFR2*- has a high degree of similarity at protein level to that found in *E. coli* sharing 22.5 % sequence identity. Analysis of the flanking genes of the other possible *gly1* genes revealed MHpara to be situated downstream of a putative *Mannheimia haemolytica* haemoglobin receptor (HmbR), along with other genes predicted to have an involvement in haem uptake (Figure 3.5). The flanking genes of the remaining genes of interest were also investigated (Figure 3.6), however, the majority of these coded for 'hypothetical' proteins, even when a BLAST search was performed with their respective sequences. Searches upstream of the MHpara gene uncovered genes for periplasmic transporter systems, including a possible PotC which is known to be involved in the periplasmic transport of putrescine and spermidine in *E. coli*.

The gene coding for a putative General control of amino-acid synthesis 5-like (Gcn5-like) acetyltransferase flanking the COI_1584 gene could have a role in transferring acetyl groups between acetyl coenzyme A and other proteins. The integrative conjugative element protein, or ICEP, labelled in the flanking genes of COI_1584 and MHA_0579 is a predicted mobile genetic element, and therefore does not provide any further indication of the possible function of the Gly1 proteins. This could suggest a similar mobile element to the *gly1* genes.

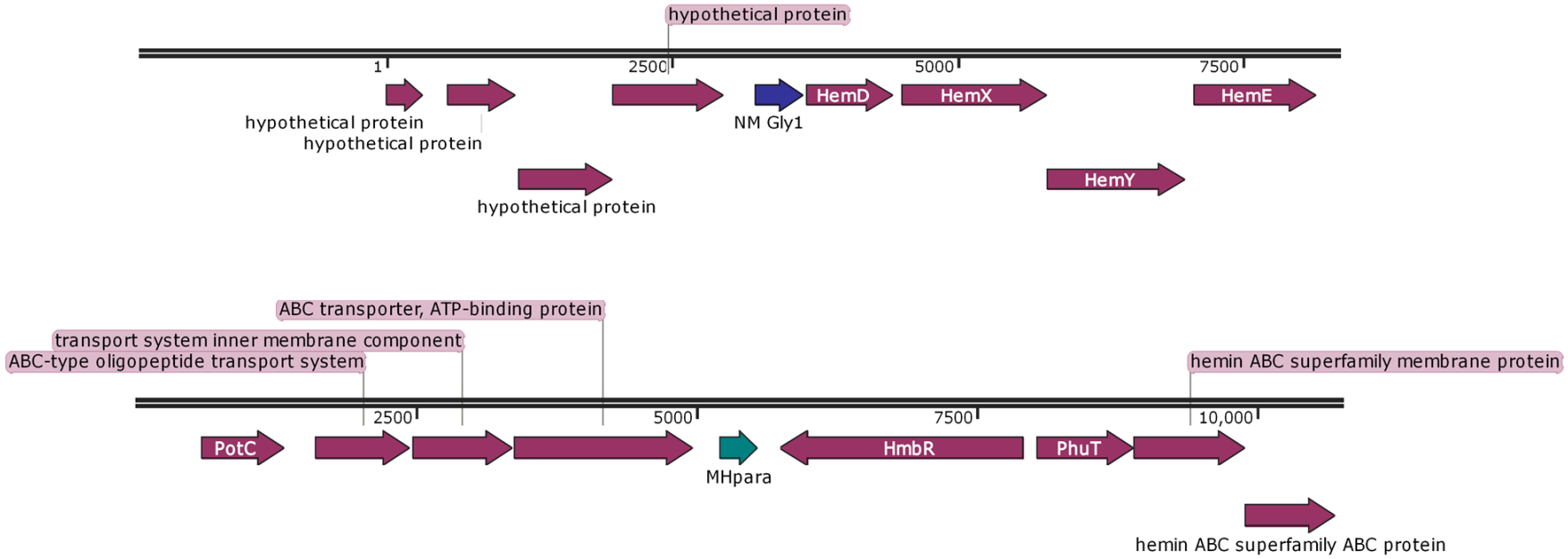


Figure 3.5. Flanking Genes of Interest.

Neisserial flanking genes upstream of the *gly1ORF1* gene are HemD (uroporphyrinogen III synthase, the enzyme responsible for uroporphyrinogen III synthesis), HemX (a suspected uroporphyrin-III C-methyltransferase), HemY (coproporphyrinogen oxidase which catalyses the formation of protoporphyrinogen III), and HemE (protoporphyrinogen oxidase, converts protoporphyrinogen III to protoporphyrin IX). Those downstream were genes coding for proteins of currently unknown function. The *MHpara* gene was flanked upstream by genes coding for proteins involved in haem uptake, and downstream by those involved in transportation. HmbR- haemoglobin outer membrane receptor, PhuT- periplasmic haem binding protein. Accession numbers as stated in Figure 3.2.

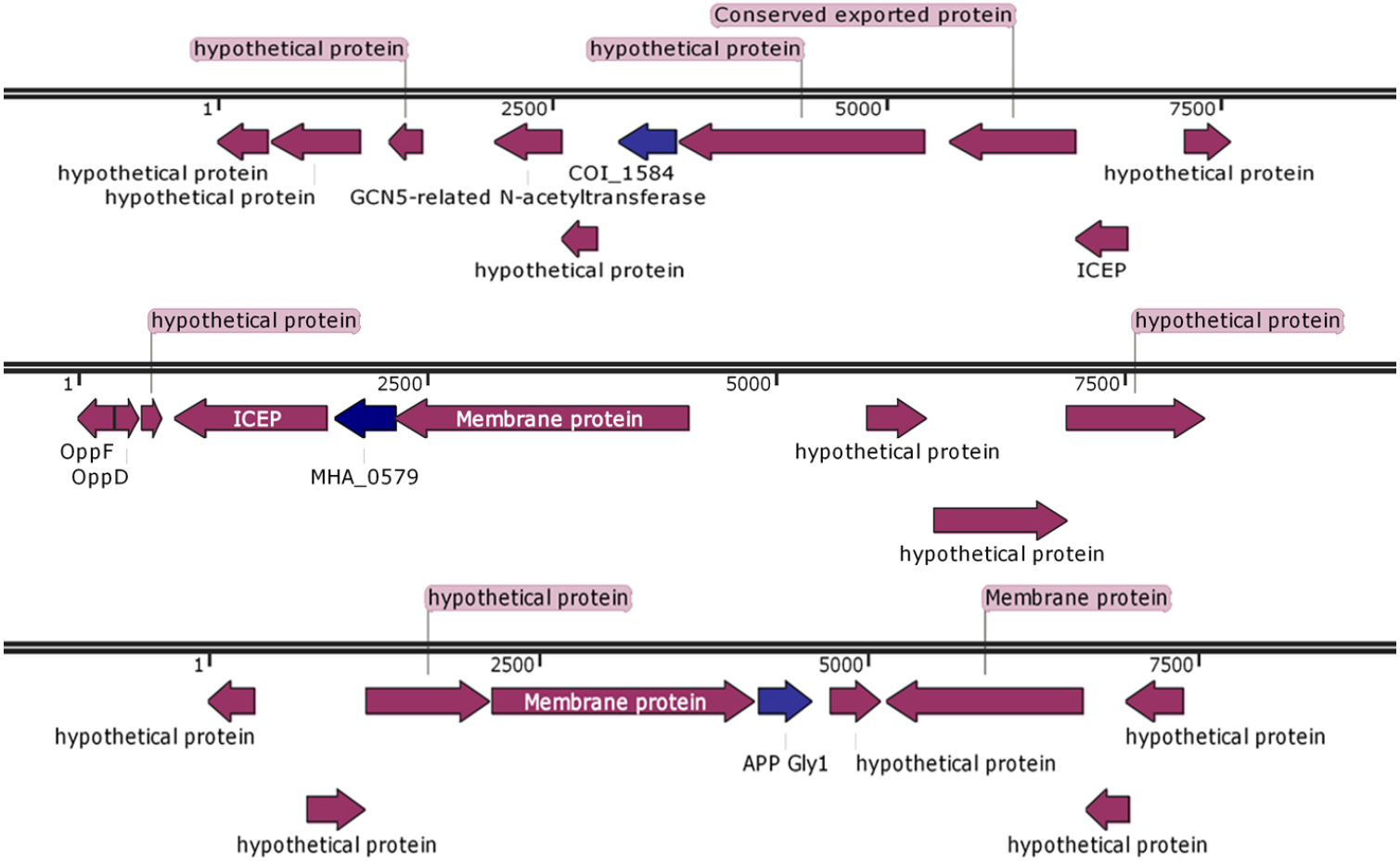


Figure 3.6 Flanking Genes of Gly1 Homologues.
Flanking genes of certain Gly1 homologues were mostly genes coding for 'hypothetical proteins'. One recurring gene is that for suspected membrane proteins of unknown function. The *S. enterica* Gly1 flanking genes are all labelled as coding for hypothetical proteins (not shown).

3.2.2 Cloning, Expression and Purification

To enable *in vitro* investigation into the structure and function of the potential Gly1 homologues identified in the aforementioned database searches, the genes coding for proteins of interest were either cloned directly from *M. haemolytica* genomic DNA (Figure 3.7) or obtained by gene synthesis (EurofinsMWG), and transformed into *E. coli* cells for over expression. Protocols for overexpression and purification were optimised for each protein (Figures 3.8- 3.10).

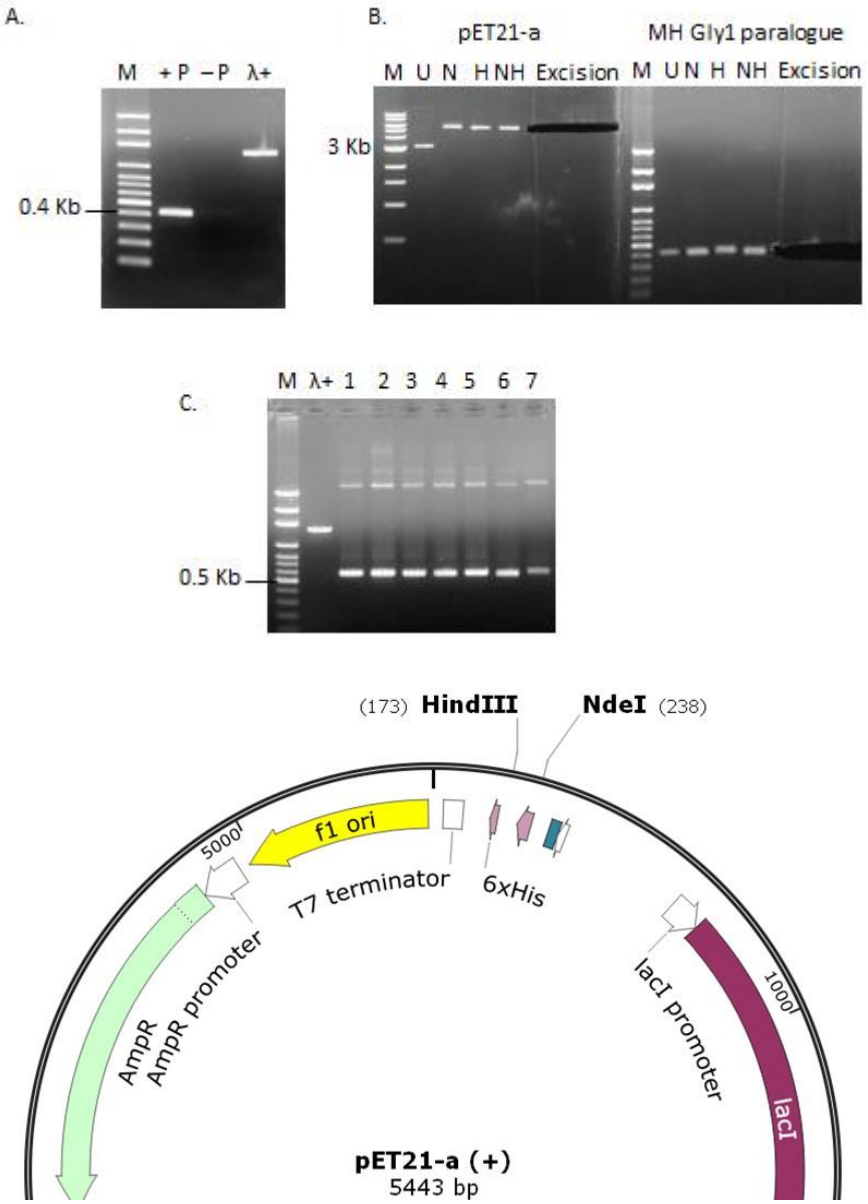


Figure 3.7. Cloning of *M. haemolytica gly1* Paralogue.

A. PCR on the *M. haemolytica gly1* paralogue gene cloned from genomic DNA (+P), reaction mixture without primers (-P) and the lambda positive control. **B.** A double digest of pET21-a and the *gly1* paralogue with NdeI and HindIII 'N' and 'H' depict single digests, and 'NH' the double digest. 'U' denotes undigested plasmid. Restriction was followed by ligation and transformation into XLIB *E. coli* cells and PCR to confirm the genes presence, results of which are shown in panel **C,** all separated by agarose gel electrophoresis. Partial pET21-a plasmid map with restriction sites denoted in the lower panel. Edited in SnapGene.

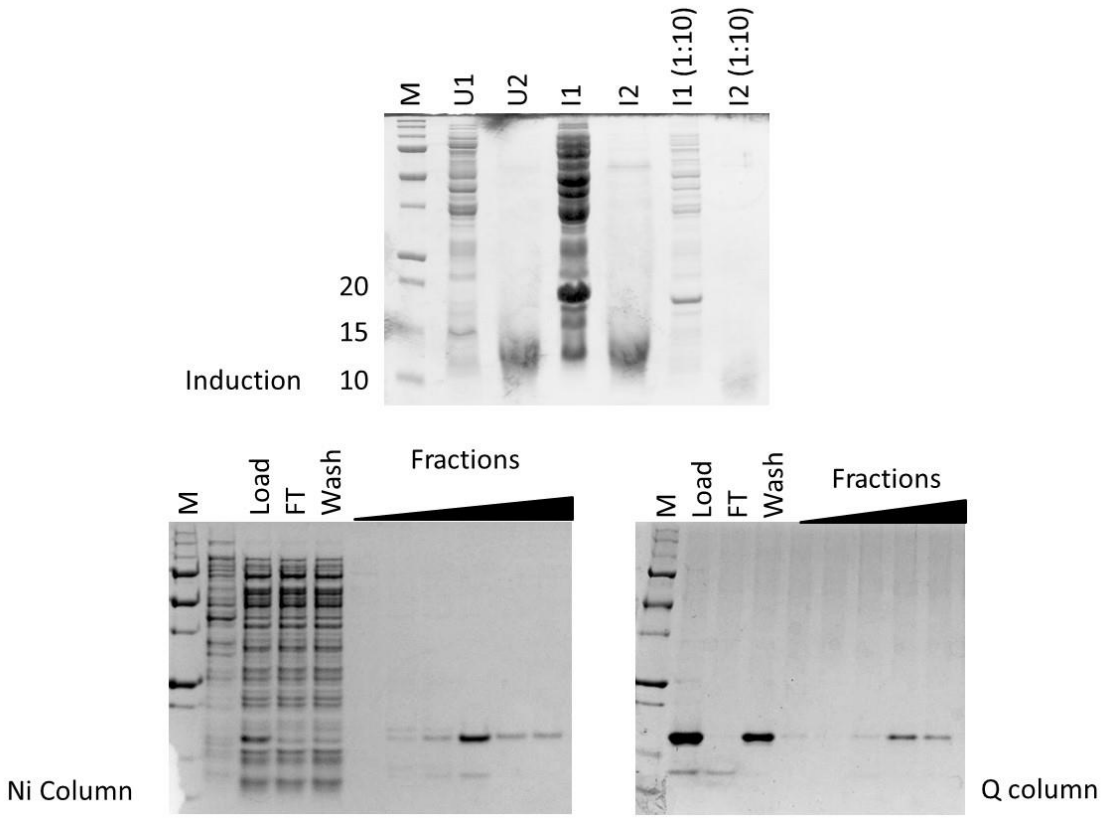


Figure 3.8. Purification of MHA_1205 C-His Gly1.

The polyhistidine tagged MHA_1205 Gly1 homologue was overexpressed in *E. coli* M72 cells. The first panel shows SDS-PAGE analysis of pre-induction cell pellet, U1, and supernatant U2, compared with the post-induction cell pellet, I1, and supernatant I2, with the band of overexpressed seen at approximately 17 kDa. The 1:10 dilution of lanes I1 and I2 clarifies the protein over-expression. The lysate was passed through a nickel chelate column (Ni), removing the majority of contaminants in the flow through (FT) and wash. A final anion exchange 'Q' column removed the residual contaminating band. MHA_0579-c-his was also purified using this protocol.

Neither the *M. haemolytica* proteins nor the APP protein were found to be secreted by *E. coli* into the cell supernatant, as shown in the SDS-PAGE gels above. The majority of MHA_1205 C-His was expressed in an insoluble form, so the purification was performed on the soluble fraction only. Due to the unknown function it would have been difficult to assess whether the protein had refolded correctly after solubilisation treatment. Some previous overexpression procedures showed MHA_1205 C-His to be secreted however this was not observed under these conditions. Batch to batch variation was observed upon purification of this protein. The elution of MHA_1205 C-His in increasing amounts in the early fractions as well as in the column wash could be explained by

insufficient resuspension of the first three elutions prior to loading on an SDS-PAGE gel for separation (Figure 3.8). MHA_0579 C-His and MHpara expression produced a good yield of approximately 2- 3 mg of soluble protein per gram of cells (Figure 3.9).

N terminal sequencing of the purified APP Gly1 protein began with the residues 'E-N-I-S-I-F', and the MHpara protein with residues 'A-G-E-Q-S-D-A', showing that these proteins had been processed as predicted by SignalP (performed by Dr. Arthur Moir, Department of Molecular Biology and Biotechnology, TUOS).

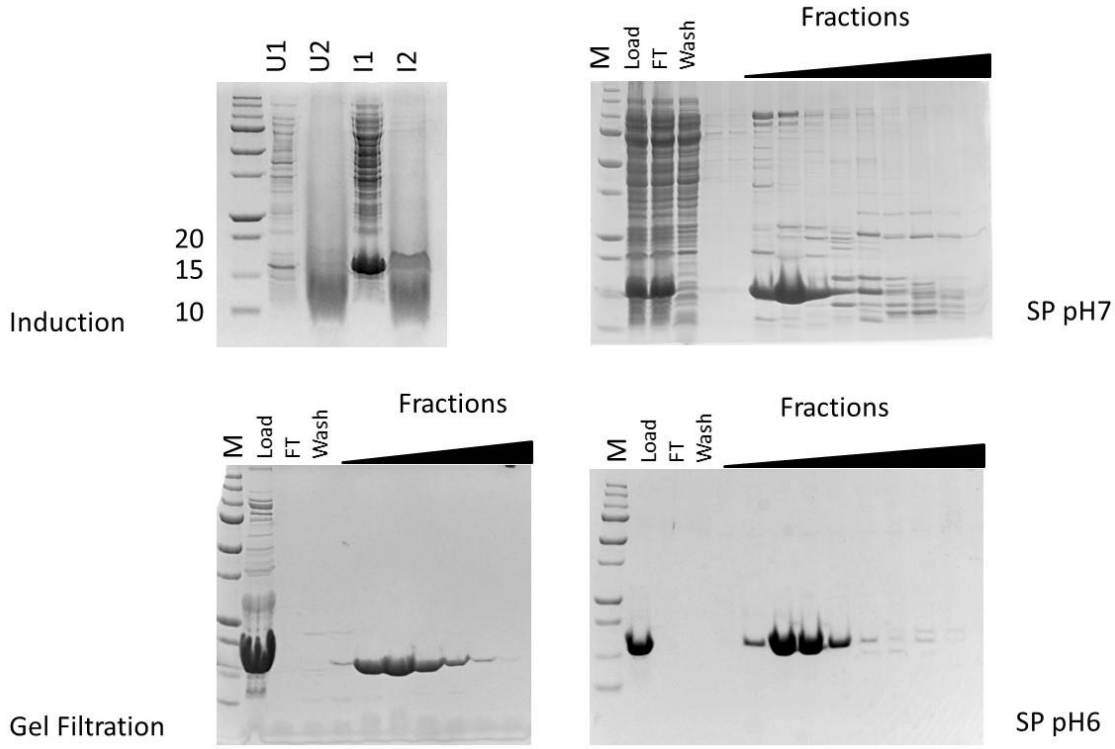


Figure 3.9. Purification of *M. haemolytica* Gly1 Paralogue.

The Gly1 paralogue, overexpressed in *E. coli* BL21 DE3 cells, was purified using anion exchange and gel filtration and analysed by SDS-PAGE. U1 shows the pre-induction cell pellet and U2 the supernatant. I1 shows the post-induction cell pellet and I2 the post-induction supernatant. The overexpressed protein is seen as a band of approximately 16 kDa. A cation exchange 'SP' column pH 7 followed by gel filtration and a final cation exchange 'SP' column pH 6 completed the purification.

PhuT was both secreted from the *E. coli* cells (Figure 3.10, Panel B) in a soluble form and found in the cellular pellet in insoluble form. The secreted protein was concentrated by diafiltration using VivaSpin columns (see Methods, section 2.6.3) and purified using a Q

column pH 8 followed by an SP column pH 6 (Figure 3.10). The insoluble protein was purified in the same manner following solubilisation, however, much of the total quantity of protein expressed remained insoluble.

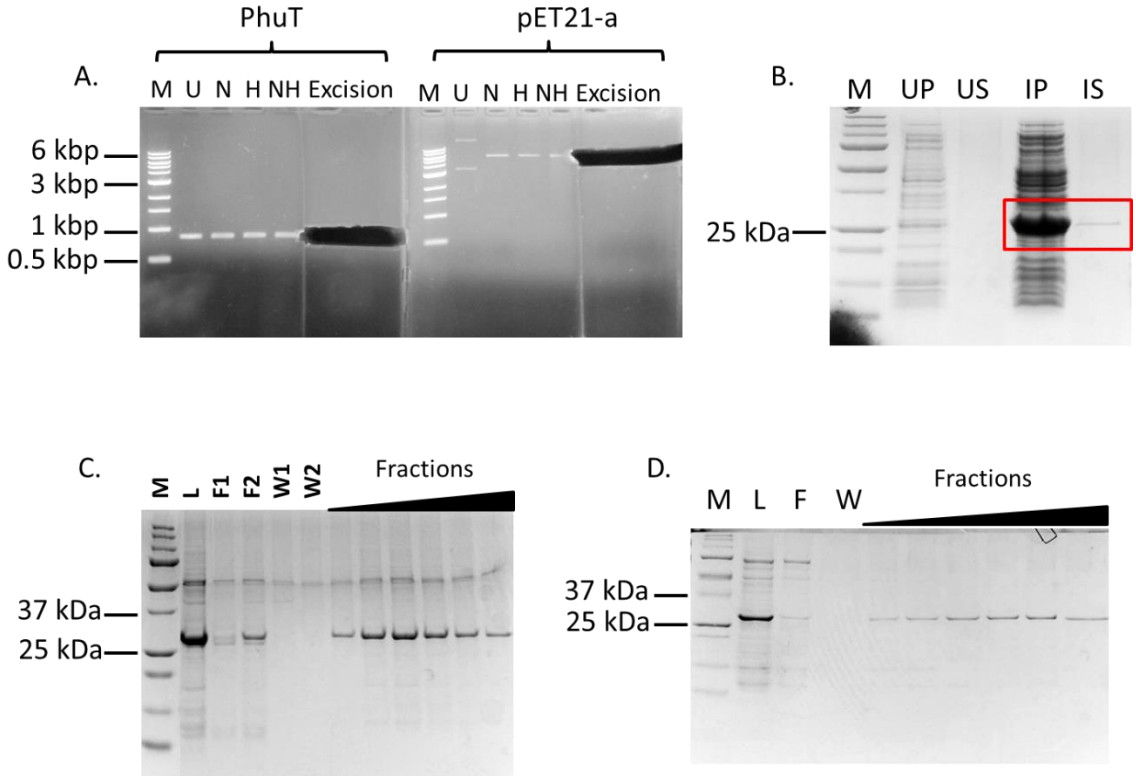


Figure 3.10. Purification of PhuT.

The periplasmic haem binding protein identified in *M. haemolytica* was cloned from *M. haemolytica* genomic DNA and inserted into the pET21-a vector- agarose gel electrophoresis of PCR products and vector fragments used for cloning PhuT (panel A). SDS-PAGE separation of the over-expression of PhuT induced by IPTG produced an insoluble protein Lane IP (induced pellet) as monitored by SDS-PAGE separation, and a soluble secreted protein from the cell-free supernatant Lane IS (Panel B) which were purified using ion exchange chromatography (Panels C and D). UP and US again uninduced pellet and supernatant, respectively.

3.2.3 Secondary Structure Analysis of Gly1 Homologues

Proteins may have similar function if the overall structure allows, regardless of whether the polypeptides share a low percentage of sequence identity, as residues with similar properties can often be interchanged. It is therefore difficult to analyse the potential homologues without information regarding the protein structure. Secondary structures can be used to identify potential homologues that have weak sequence identity, and protein tertiary structure may enable identification of binding pockets that may be more reliant on charge than amino acid sequence. Structural folds can bring together residues that are important in ligand binding but distant in terms of the primary structure.

The secondary structure of the neisserial proteins was discovered to contain eight antiparallel beta strands separated by a small section of alpha helix. To compare these neisserial Gly1 proteins to the potential new Gly1 homologues, the secondary structures of the proteins of interest were initially assessed using online prediction software, SABLE (Adamczak et al., 2005), followed by analysis by circular dichroism (CD) on the proteins available. In the structure and activity assays, MHA_1205 C-His was used to represent the most similar *M. haemolytica* Gly1 homologues with just one or two amino acids differing between the sequences.

Analysis with the SABLE server estimated MHA_0579 and the other *M. haemolytica* homologues to have 9 - 11 beta strands contributing to approximately 34% of each structure, a high proportion of coils/ turns/ disordered regions making up 50% and a small amount of helical structure between 13 and 16% of the protein (Figure 3.11). Other species Gly1 homologues were also predicted to be mainly beta in composition (Figure 3.12). The paralogue, however, was predicted to have 8 regions of beta strands which was supported by the subsequent determination of tertiary structure (Figure 3.15 and 3.16). PhuT was predicted to be formed of 47% alpha helices, as suspected from its predicted function based on other periplasmic haem binding proteins.

The Phyre2 server secondary structure prediction largely agreed with that created by SABLE, other than a predicted beta strand at around position 110 rather than the SABLE predicted alpha helix in mh0603, MHA_1205 and COI_1584. The SABLE server predicted both alpha helical and beta strands at the IKKAV region, which differed between

proteins, whereas Phyre2 suggested all alpha at these points. Additionally an alpha helical region was predicted by Phyre2 at the C terminal of MHA_0579 where SABLE predicted beta strands. Phyre2 predicted that the 5th and 6th beta strands in SARI were just one continuous beta strand and an additional short beta strand was predicted of 'ESR', at around position 100. Secondary structure analysis was performed with the predicted signal peptides present at the N terminal region, and the first beta strands predicted in MHA_1205, mh0603 and COI_1584 reside upstream of the predicted cleavage site.

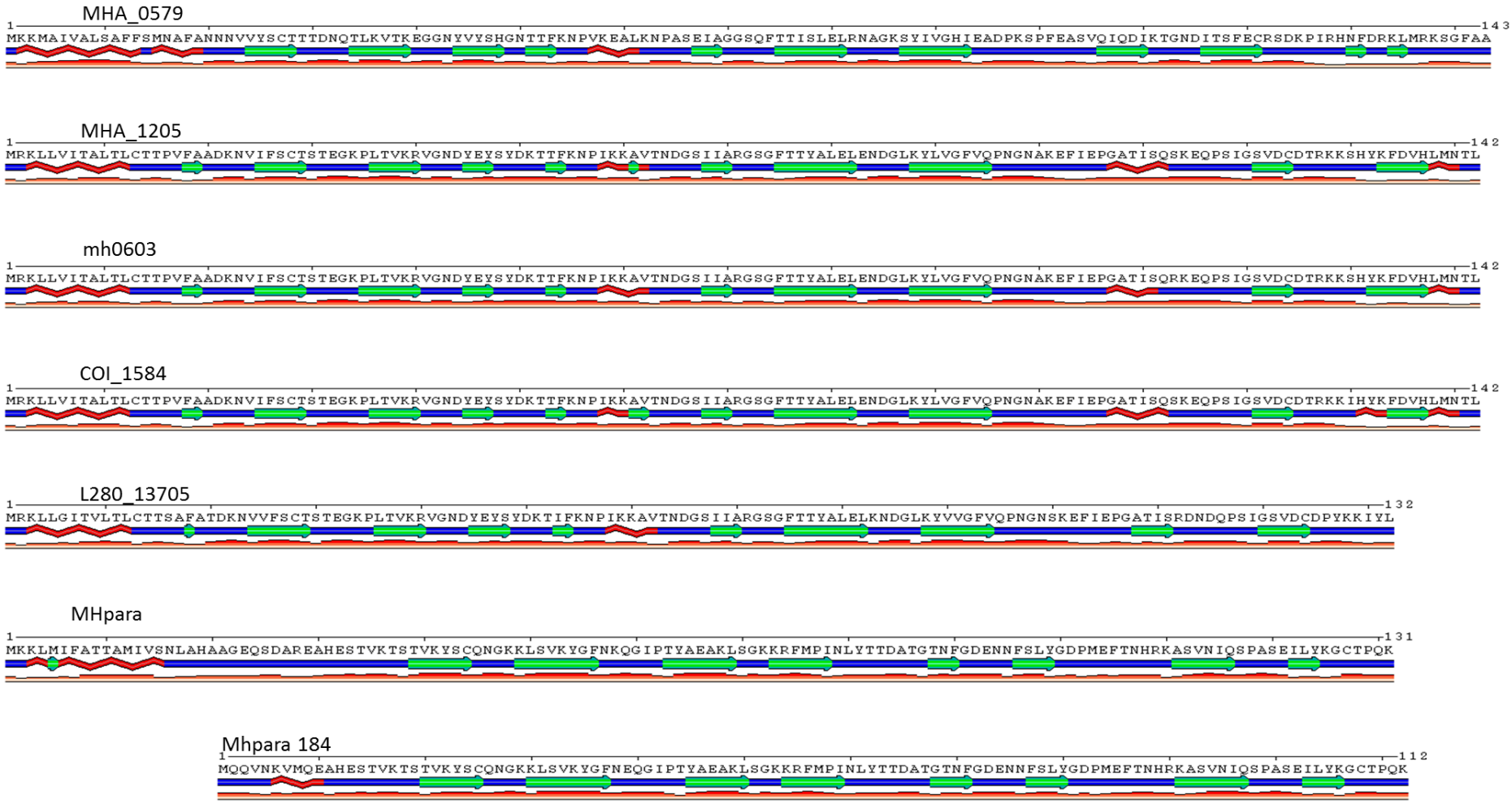


Figure 3.11. Secondary Structure Prediction of *M. haemolytica* Proteins of Interest Using SABLE Red zigzags show predicted helical structures, green arrows show predicted beta structures. The red graph on the lower line indicates confidence in secondary structure prediction- high bars equates to high confidence.

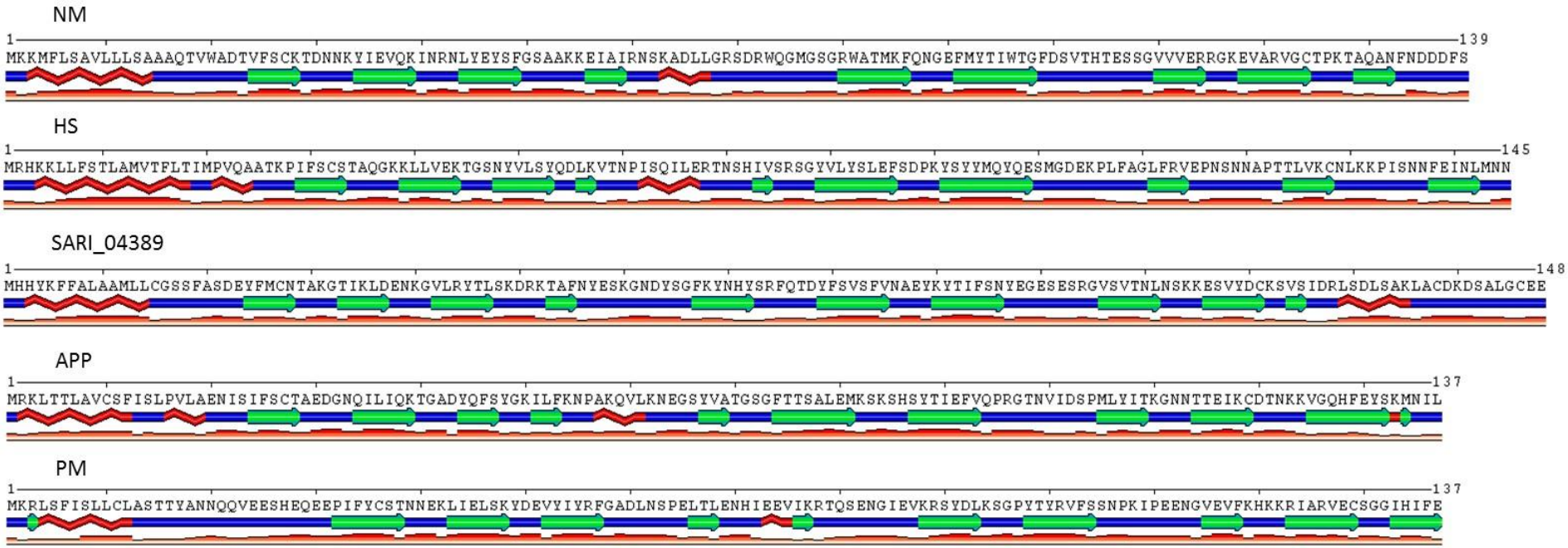


Figure 3.12. Secondary Structure Prediction of Other Gram-negative Bacterial Proteins of Interest Using SABLE

Red zigzags show predicted helical structures, green arrows show predicted beta structures. The red graph on the lower line indicates confidence in secondary structure prediction- higher bars equate to higher confidence. NM depicts the 'original' *N. meningitidis* Gly1 protein complete with eight beta strands. HS is the *H. somni* predicted Gly1 and PM is *P. multocida*.

Once the proteins of interest had been expressed, circular dichroism was used to further analyse the protein secondary structure (see figure 3.13) to allow comparison with the neisserial Gly1 protein. Based on the commonly observed CD spectra of proteins, those with mainly antiparallel beta secondary structure generally have a negative peak at 216 nm and a positive peak at 195 nm (Greenfield, 2006). MHA_1205 C-His, MHA_0579 C-His and APP all had spectra resembling those of proteins consisting of a high percentage of beta strands. This is supported by the lowest point of the spectra, which was observed at 216 nm for MHA_1205 C-His and at 218 nm for MHA_0579 C-His. With APP, the negative peak occurred at 219 nm. All three had positive peaks at between 193 and 199 nm. These positive peaks were measured at atmospheric pressure so must be analysed with caution, as at wavelengths below 200 nm air can absorb UV.

PhuT had a spectrum indicating a mainly alpha-helical secondary structure, which supports the predicted secondary structure seen with SABLE analysis. Alpha helical proteins generally have a double negative peak at 222 nm and 208 nm with a positive peak at 192 nm (Greenfield, 2006). In this instance the negative peaks were observed at 217 nm and 208 nm, and the positive peak at 195 nm.

The spectrum of MHpara looked to be mainly beta but had a slight second negative dip suggesting that another structure was possibly affecting the shape. The protein gave a positive peak at 233 nm, a negative peak at 211 nm and a further positive peak at 195 nm. SARI C-His, also predicted to have a mainly beta secondary structure, gave a similar double negative peak spectrum resembling that of helical reference proteins. The negative peaks were observed at 216 nm and 201 nm and a positive peak at 194 nm.

To further analyse the data generated, DichroWeb (Whitmore and Wallace, 2008) was used, with the reference set SMP180 (Abdul-Gader et al., 2011), as this provides a broad coverage of proteins, and analysis programme CONTIN (van Stokkum et al., 1990). The results for MHA_0579 C-His, MHA_1205 C-His, SARI C-His and MHpara agreed with the SABLE predicted percentages of alpha and beta forming the secondary structures. The PhuT spectral analysis on average suggested that the protein was 19% alpha, 27% beta and 54% coils/ turns/ disordered protein. This did not agree with the predicted secondary structures generated by the SABLE and Phyre2 servers.

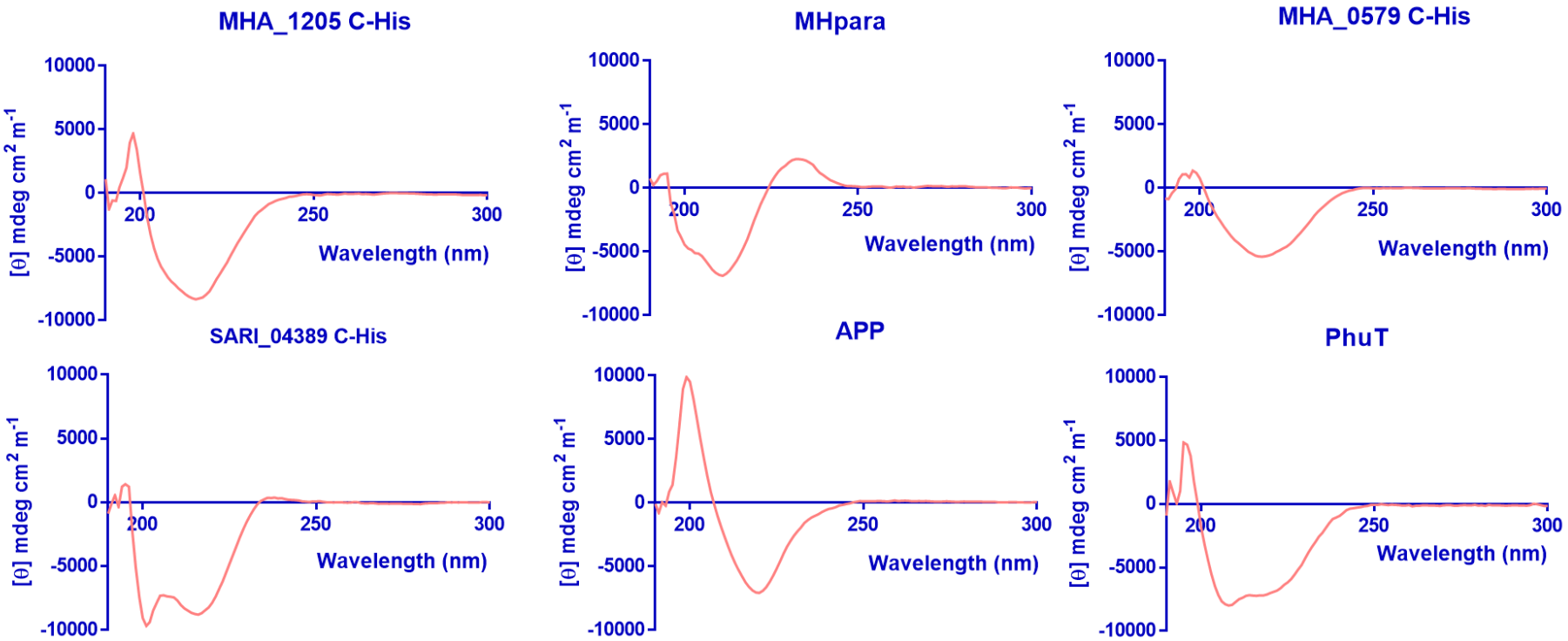


Figure 3.13. Secondary Structure Analysis Using CD Spectra.
The CD spectra of Gly1 homologues showed curves characteristic of antiparallel beta strands (MHA_1205 C-His, MHA_0579 C-His and APP) or containing some alpha helical structures (MHpara and SARI C-His). PhuT also had a spectrum suggesting alpha helical structures. N=3.

3.2.4 Protein Tertiary Structure

To investigate further structural similarities between suspected Gly1 homologues, X-ray crystallography was performed on purified protein samples. Solved structures could also provide clues to putative ligand binding pockets. Although crystals of MHA_1205 C-His and SARI C-His were produced using the hanging drop technique, they did not diffract well enough to allow structural analysis. MHA_0579 C-His crystals grown in PACT D3, which is 0.1 M MMT buffer pH 9 and 25% (w/v) PEG 1500 diffracted to 2.1 Å but the structure has not been solved to date. Additional attempts to solve this were made by soaking the crystal in hemin, however, further optimisation is required to produce results. Figure 3.14 shows the protein crystals.

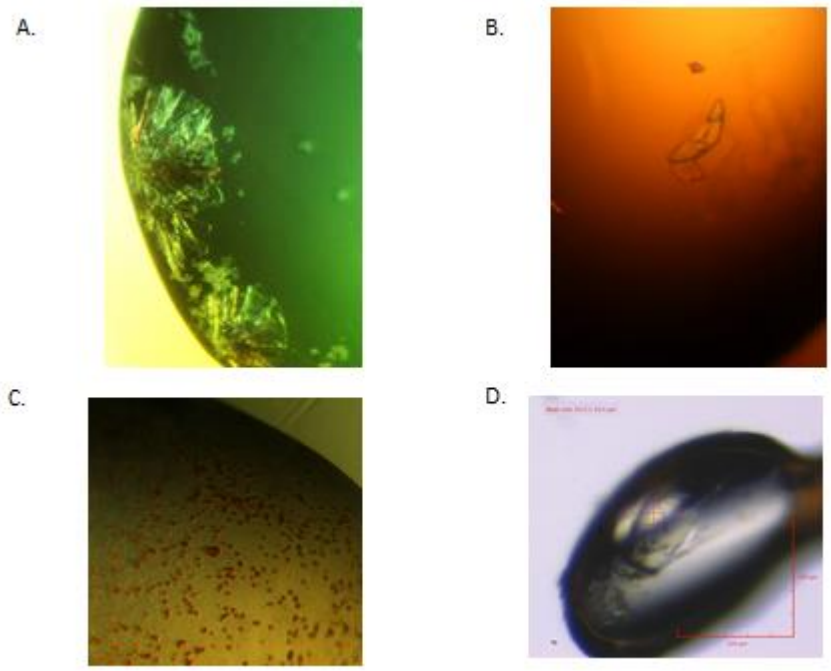


Figure 3.14. Crystals of Gly1 Homologues.

Structures of the crystallised proteins seen above have not yet been solved. **A.** SARI C-His in 0.2 M potassium phosphate and 20% PEG3350. 10 weeks incubation at 17°C (by Richard Salmon, TUOS) **B.** MHA_0579 C-His PACT D3 (0.1 M MMT buffer pH 6 and 25% (w/v) PEG 1500) **C.** and **D.** MHA_0579 C-His in PACT D3 (0.1 M MMT buffer pH 6 and 25% (w/v) PEG 1500) after hemin soaking respectively (B, C and D by Jason Wilson, TUOS). A, B and C at 68x magnification. Red scale bar in D is 100 µM.

The *M. haemolytica* Gly1 paralogue was crystallised in JCSG+ A5 which contains 0.2 M magnesium formate and 20% PEG3350 to 1.37 Å (see Figure 3.16). No structure was solved for the other Gly1 homologues. The MHpara protein was crystallised in dimer formation as shown in the structure solved by Jason Wilson (Figure 3.15). The crystal structure of the *M. Haemolytica* Gly1 paralogue shown in Figure 3.16 has the final cysteine highlighted and shows the antiparallel beta strands forming a barrel like structure. The paralogue protein comprises an YSC region rather than the FSC region used for identification initially, and the FSC can be seen in the *N. gonorrhoea* structure (Figure 3.4). alongside the neisserial structures generated by Magda Weirzbicka/ Jason Wilson (Department of Molecular Biology and Biotechnology, TUOS, 2014) and Arvidson *et al* (Arvidson et al., 2003).

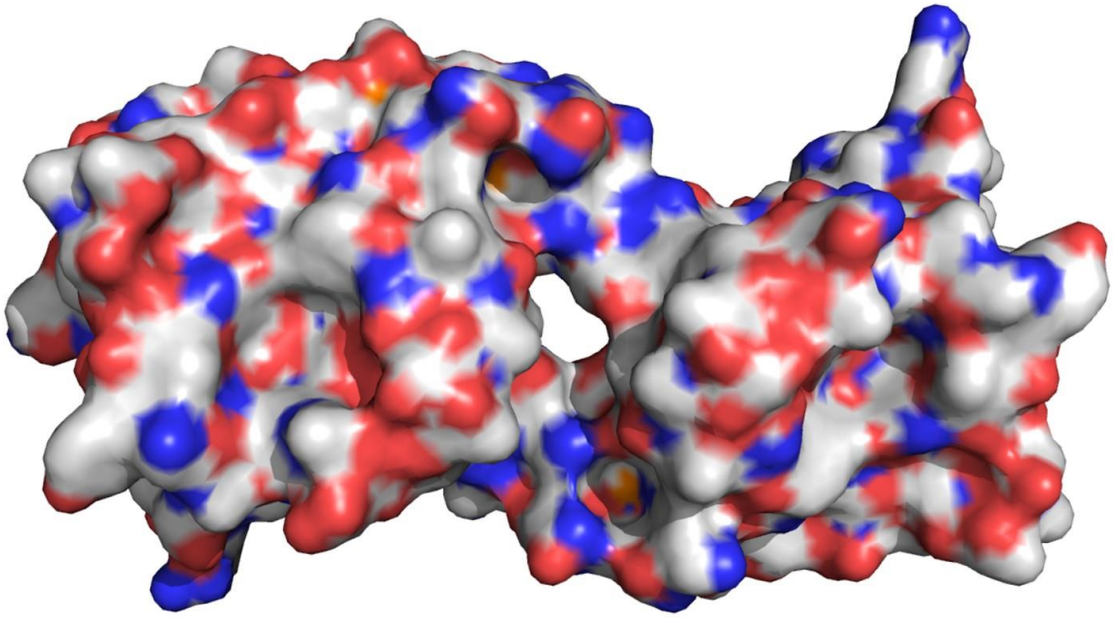


Figure 3.15 *M. haemolytica* Paralogue Crystal Dimer

The protein was crystallised in dimer formation as seen above. It is not known whether this is the natural state of the protein or an artefact of the crystallisation procedure. Cysteine residues in orange, carbon atoms white, oxygen atoms red and nitrogen atoms blue. Rendered in PyMol.

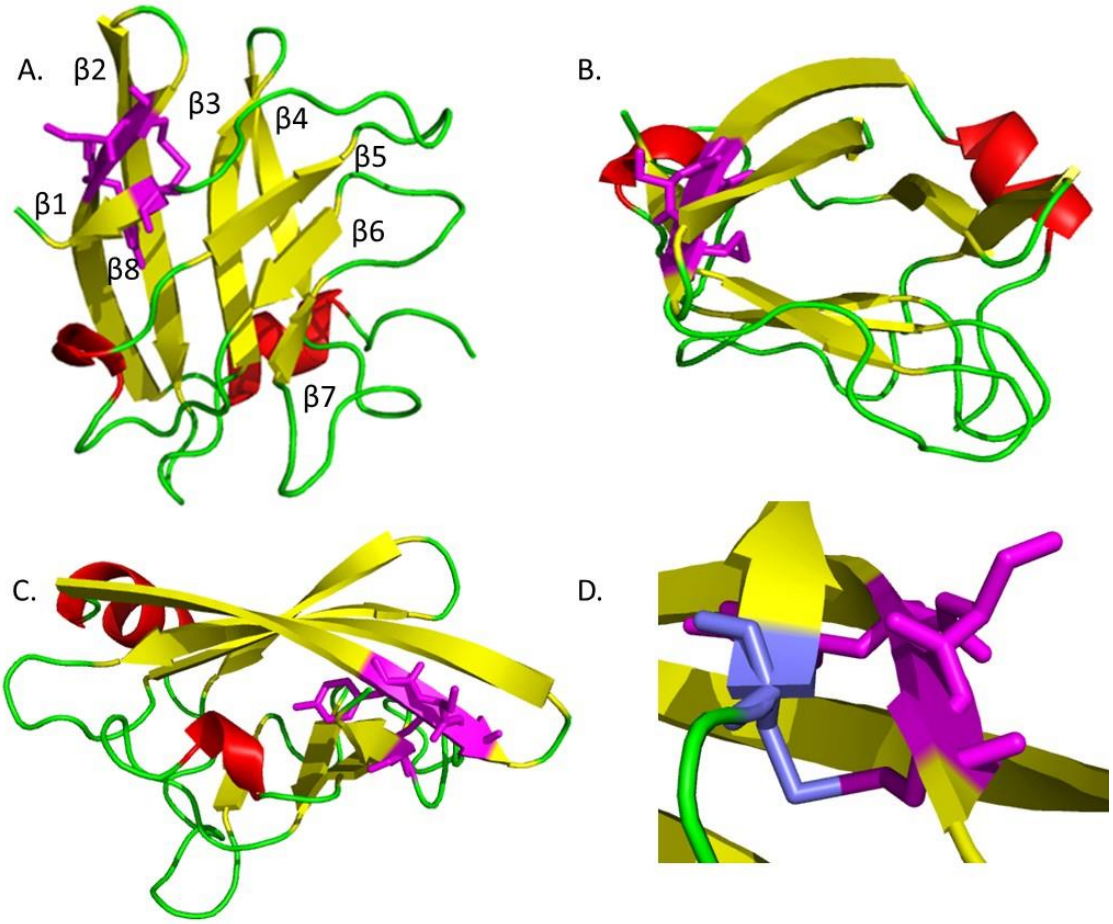


Figure 3.16. ***M. haemolytica*** **Gly1 Paralogue Crystal Structure.**

Panel **A** shows the antiparallel beta sheet structure of MHpara with a minority made up of alpha helix viewed from the *x* axis. Beta strands numbered from the N- terminal. Panel **B** and Panel **C** show views from the *y* and *z* axes, and how the two highlighted regions interact with one another. The FMC and C terminal cysteine are highlighted in magenta. Panel **D** shows the disulphide bridge formed between the two cysteines, with the C terminal cysteine this time highlighted in blue. Structure rendered in PyMol.

An overlay of the crystal structures of *N. gonorrhoeae* Gly1ORF1 and the MHpara protein (Figure 3.17) highlights the similarities despite the low protein sequence identity of 16.4% at the amino acid level.

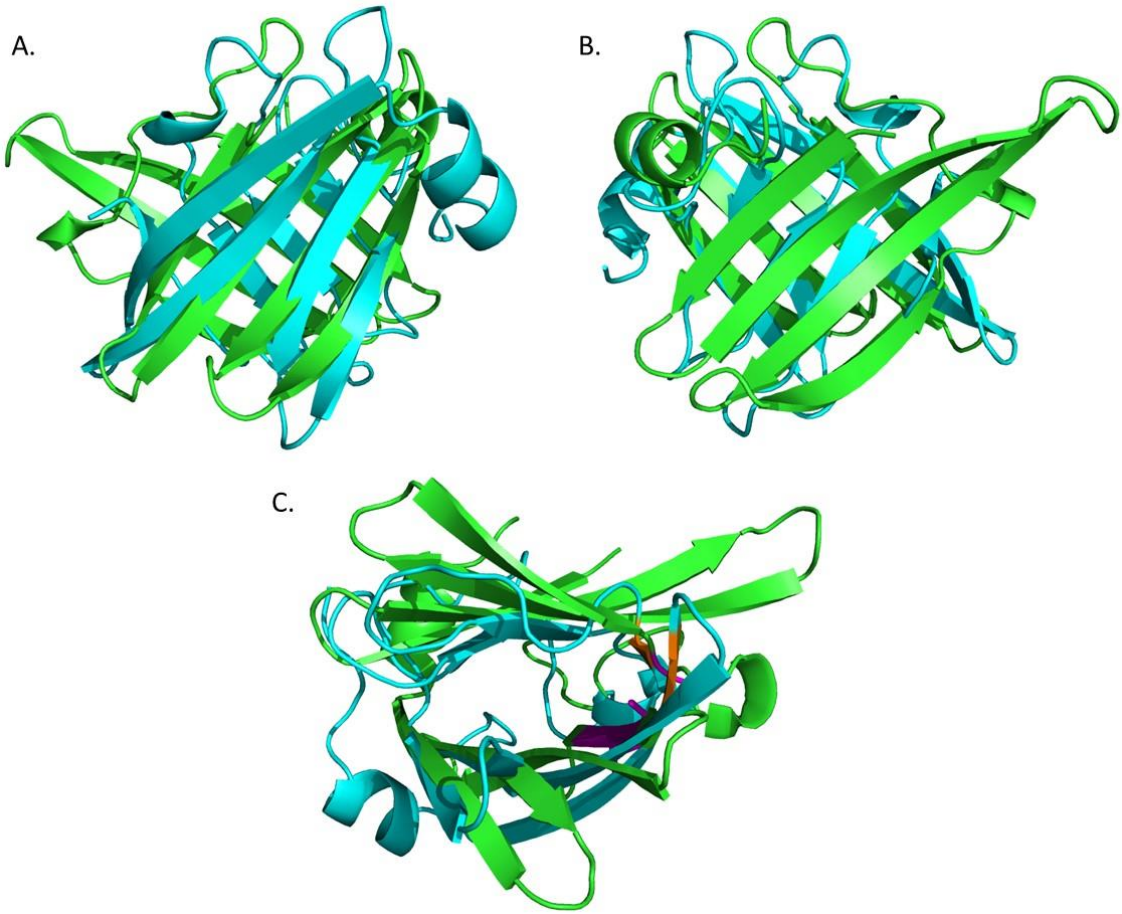


Figure 3.17 Overlay of *M. haemolytica* Paralogue and *N. gonorrhoea* Gly1ORF1 Structures.

N. gonorrhoeae Gly1ORF1 in green and MHpara in cyan. Three orthogonal views of the overlayed structures with the FSC and C of *N. gonorrhoeae* Gly1 in purple and YMC and C of MHpara in orange in panel **B**.

Phyre2 was used to predict the tertiary structure of the proteins of interest that were not solved during the course of the project (Figure 3.18). Proteins had their signal peptides removed prior to input, as those purified for this study lacking the polyhistidine tag underwent N terminal sequencing, and this showed they had been processed.

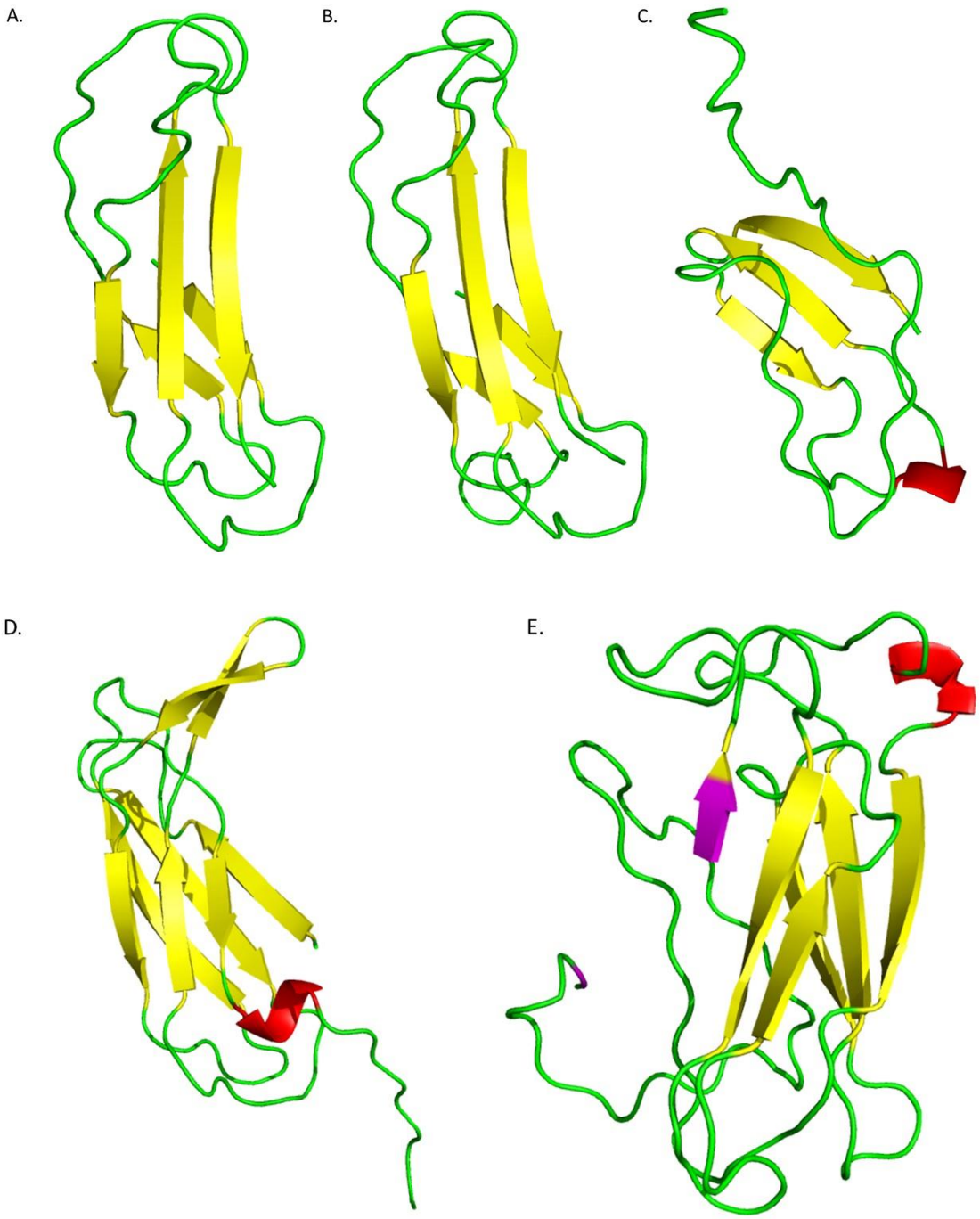


Figure 3.18. Predicted Tertiary Structure of Gly1 Homologues.

A. MHA_0579 confidence of 48.9% over 45% coverage **B.** MHA_1205 confidence of 44.8% over 46% coverage, both based on part of structure 2XZZ **C.** *P. multocida* confidence of 27.8% over 44% coverage, component of structure 1F1S **D.** APP 56.4% confidence and 66% coverage also based on human transglutaminase 1 beta barrel domain 2XZZ **E.** SARI 25.5% confidence and 56% coverage based on galactose binding protein 1DLC. Structure predicted by Phyre2 server and model edited in Pymol. PDB reference numbers used for identification.

The confidence in the predicted tertiary structure of the *M. haemolytica* Gly1 homologues was very low. The models only covered a small percentage of the amino acid sequences leaving a lot of missing information (Figure 3.19). They do, however, go some way to support the predicted secondary structures as they are mainly composed of beta strands.

MHA_1205:
MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYD**KTTFKNPIKKAVTNDGSIIARGSGFTTYAL**
ELENDGLKYLVGFVQPNGNAKEFIEPGATISQSKEQPSIGSVDCDTRKKSHYKFDVHLMNTL
MHA_0579:
MKKMAIVALSAFFSMNAFANNNVVYSCTTTDNQTLKVTKEGGNYVYSHG**NTTFKNPVKEALKNPASEIAGGSQFTTISLE**
LRNAGKSYIVGHIEADPKSPFEASVQIQDIKTGNDITSFECRSDKPIRHNFDRKLMRKSGFAA

Figure 3.19 Areas Covered by Tertiary Structure Prediction of *M. haemolytica* Proteins.

Sequences used to create predicted tertiary structures with the Phyre2 server. The sections highlighted in bold are those used for tertiary structure prediction. The key regions (FSC, YEYS, C) are not included in these predictions.

This also highlights the novel structure of this family of proteins as there was nothing appropriate to model them against in the database used by the Phyre2 server. It is interesting to note that MHA_1205, MHA_0579 and APP were all modelled on the same crystal structure- PDB reference 2XZZ, the crystal structure of the human transglutaminase 1 beta-barrel domain, which is an eighth of the 817 a.a. protein. This is involved in protein cross linking such as the clotting factor XIII. The active site residues in the human clotting factor XIII form a triad of Cys-His-Asp. These three residues are not adjacent to each other in the amino acid sequence but are brought together in the tertiary structure (Yee et al., 1994). The proteins modelled on this section of transglutaminase contain the three residues that form the active site, however, the histidines were not included in the region for which the partial predicted tertiary structure was generated so we cannot be sure if they are located in positions creating an active site.

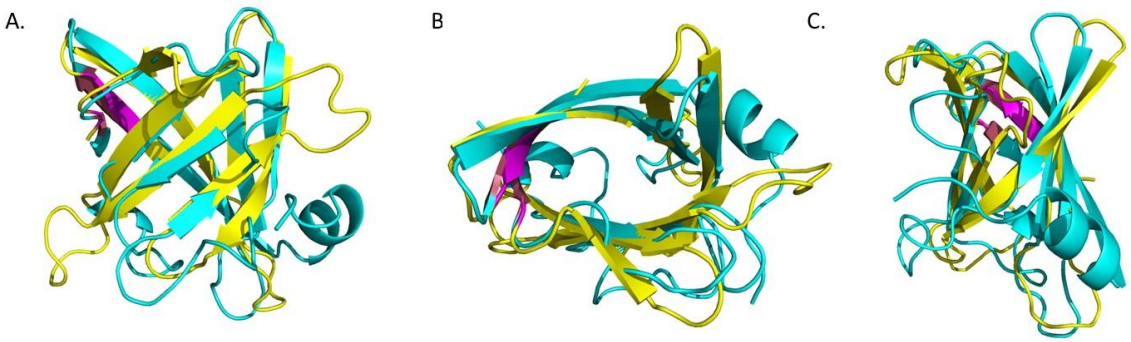


Figure 3.20. Predicted Tertiary Structure of MHpara.

Panels **A, B,** and **C** show three orthogonal views of the MHpara model (yellow) aligned with the structure solved by Jason Wilson (cyan). The predicted structure was based on the template of protein 3OE3 from the PDB site, the crystal structure of PliC-st with 96.8% confidence of 60% coverage.

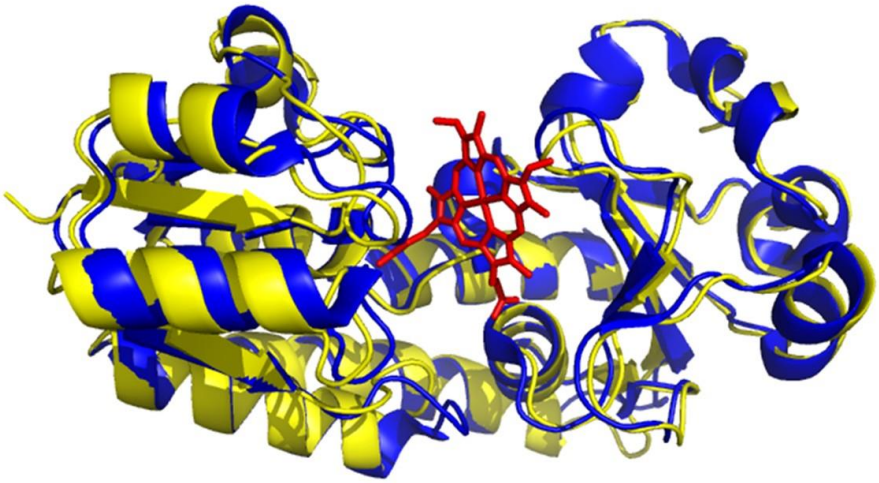


Figure 3.21 Predicted Tertiary Structure of PhuT.

ShuT from *Shigella dysenteriae* (blue) aligned with the predicted PhuT structure of *M. haemolytica* (yellow) based on 2R97, a *P. aeruginosa* haem PBP. The haem molecule in red is shown as seen in the crystal structure of the ShuT protein (Ho et al., 2007). This crystal structure was solved as a four protein complex. Proteins modelled in Phyre2 and edited in PyMol.

Confidence in the MHpara tertiary structure model was 96.8% over a coverage of 60% of the amino acid sequence. This can be seen in Figure 3.20, where the solved structure is aligned with the predicted structure. The general shape and secondary structure of

the predicted MHpara aligns to the actual solved structure, although the beta strands are shorter in reality and the helical regions are missing in the predicted structure. The pore is also partially occluded by disordered loops, which could potentially be flexible or involved in protein interactions as has been described for a number of hemoproteins. The model is based on periplasmic lysozyme inhibitor of c-type2 lysozyme PliC from *Salmonella typhimurium*, which as the name states is a lysozyme inhibiter, which was shown to dimerise (Leysen et al., 2011). The amino acid sequences share only 18.42% identity. PhuT models had 100% confidence over 90% coverage, suggesting that these predictions give a good indication of the likely protein structure. The predicted structure of PhuT was based on a haem PBP from *P. aeruginosa*, however its close alignment with ShuT of *S. dysenteriae* enabled an insight into the possible haem binding cleft for this protein (Figure 3.21).

3.2.5 Predicted Cellular Locations of Gly1-like Proteins

The location of the proteins within the cell can indicate whether the putative function of the protein is feasible. Initially, an online programme called PSORT (Nakai and Horton, 1999), a cellular location prediction tool, was used. This analysed the probability of the proteins of interest being in either the outer membrane, the periplasm (as is hypothesised for PhuT), secreted or in other regions of the cell. Table 3.4 shows a summary of the predicted locations.

As previously mentioned, the Gly1 family of proteins share the feature of an N- terminal signal peptide which would direct the protein into at least the periplasmic space. However, it is currently unknown as to whether this is where the different proteins remain or if the function of the proteins require embedding in the outer membrane or secretion. We know PhuT is predicted to be a periplasmic haem binding protein based on the similarities to other PBP's such as ShuT, and the high degree of certainty shown by this programme supports this that it is indeed a periplasmic protein.

Protein	Periplasm	Outer Membrane
	Certainty	
MHA_0579	0.365	**0.826**
MHA_1205	**0.524**	0.252
MH0603	**0.655**	0.252
COI_1584	**0.655**	0.252
L280_13705	**0.699**	0.183
MHpara	0.229	0.177
APP	0.104	**0.775**
NM Gly1	0.268	**0.934**
NG Gly1	0.228	**0.832**
PhuT	**0.929**	0.211
HmbR	0.131	**0.922**
SARI	0.300	**0.937**

Table 3.4. PSORT Results for Proteins of Interest.

The closer the number to one, the higher the likelihood of the protein being targeted to that location based on signal peptide presence, and basic and hydrophobic regions. Those in bold are the higher numbers giving the estimated cellular location.

The neisserial Gly1 proteins are all predicted to be localised to the outer membrane of the bacteria, as are the APP, SARI and MHA_0579 Gly1's. Investigations by Sathyamurthy *et al* did not identify any OM protein motifs in the *N. meningitidis* Gly1 (Sathyamurthy, 2011). The neisserial Gly1 in the Sayers lab was, however, believed to be present on the outer membrane as the Gly1 gene knock-out bacteria was less susceptible to complement mediated bacterial killing than the wild type, and further analysis showed the Gly1 present on OM blebs (Arvidson et al., 1999, Weirzbicka, 2014, Sathyamurthy, 2011). This again could support the evidence collected so far that the proteins of interest are involved in different cellular functions. Also included is the *M. haemolytica* HmbR, a known outer membrane receptor that was predicted by the PSORT software as being in the outer membrane, as expected. The MHpara protein was not conclusively predicted to reside in either of these locations, while the results suggest that MHA_1205 and its highly similar *M. haemolytica* Gly1s are in the periplasm. Further studies would be required to state this with any confidence.

3.3 Discussion

3.3.1 Identification of Potential Homologues

The method of homologue identification was inclusive, as previously stated, with differences between the key FSC motifs being accepted, such as FMC and YSC. Tyrosine (Y) and phenylalanine (F) are both aromatic amino acids, only differing by the hydroxyl group present on the benzene ring of tyrosine. Tyrosine is more reactive to non- protein atoms than phenylalanine as a result of this and is also more soluble. Both are often buried within the protein due to their hydrophobic nature. Serine (S) and methionine (M) are both aliphatic. Methionine is hydrophobic, and again is often found buried within the protein. Serine however, is polar, and can be found in many positions within the protein. The initial drawback in searching for homologous proteins was that using online databases was hampered by the limited availability of genome sequences for strains of *M. haemolytica* when the project began. Although more have been added since, the majority currently available belong to one of three serotypes. An additional issue is the annotation of the putative open reading frames as predicted proteins which can be misinterpreted based on previous sequence annotations. The shorter MHpara proteins annotated in error in certain *M. haemolytica* genome sequences shows it is worth ensuring the hypothetical proteins presented are true to the DNA sequence.

Proteins identical to APP Gly1 were also recently found in strains of *A. equii* and *A. suis*, which could be a result of input error in the database, contamination of samples, or horizontal gene transfer. These strains with the 'multispecies' protein were added to the database in late 2014 and March 2015, after the completion of many structure-function experiments with the synthetic APP protein. The possibility of *A. pleuropneumoniae* culture contamination is greater with *A. suis* as it is also a cause of porcine disease, which is less pathogenic and often misdiagnosed as *A. pleuropneumoniae*. It is also a pathogen found in the porcine lung and could have been co-isolated. This may also provide a reason for the APP Gly1 protein only being present in one database strain of *A. pleuropneumoniae* if it is not truly native to this species.

An additional search revealed the presence of an *A. pleuropneumoniae* protein with low levels of sequence identity to Gly1 homologues (between 15- 20% amino acid identity). Further analysis showed that although the protein had an N terminal signal peptide, it did not contain the FSC or YEYS amino acid motifs or similar variants of these. Secondary and tertiary structure prediction was performed and this showed 54% alpha helical structures when analysed with Phyre2 although SABLE predicted mostly disordered or random coil with just 16% alpha helix and 20% beta strands (see figure 3.22).

The predicted tertiary structure was based on only 7% of the amino acid sequence which showed an alpha helical region (not shown). The PSORT cellular location prediction tool could not estimate where the protein exists in the bacterial cell with confidence. It is possible based on this new information that *A. pleuropneumoniae* does not contain a Gly1-like protein, and that the protein 'APP' used in the experiments in this thesis belongs to another *Actinobacillus* species.

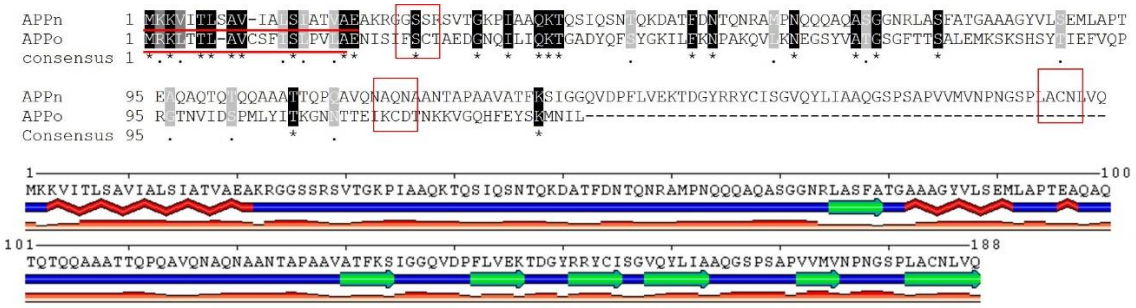


Figure 3.22 *A. pleuropneumoniae* Protein Dissimilar to Gly1 Homologues.

The signal peptides are underlined and the FSC region of the APP protein described throughout (APPo) is highlighted in the first box. The latter two boxes highlight the C terminal cysteines respectively. APPn accession number WP_005604137.1.

S. enterica arizonae (IIIa) is the only salmonella species containing a Gly1 homologue candidate to date. This subspecies was found to be the most distant relative when comparing salmonella species (Chan et al., 2003), which could provide an explanation for this. There is a protein with low level sequence identity (13.9%) to this *S. enterica arizonae* Gly1 in *S. enterica pomona.* While this protein does have an N terminal signal peptide of 26 amino acids in length, it does not possess the FSC or C terminal cysteine motifs (or variations) that we used to identify potential homologues (Figure 3.23).

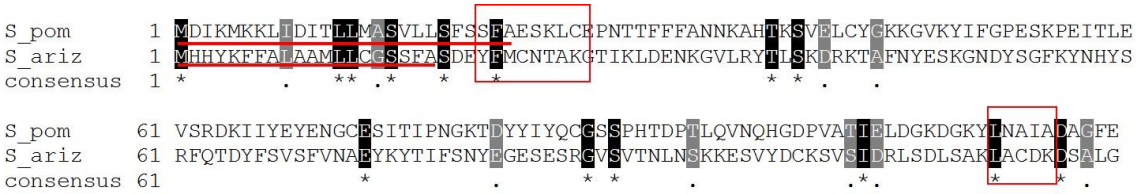


Figure 3.23 *Salmonella enterica* Protein Dissimilar to Gly1 Homologues.

Signal peptides are underlined in red, FMC region of *S. enterica arizonae* and C of *S. enterica pomona* (accession number WP_000148339.1) in the first box and the C terminal cysteine of *S. enterica arizonae* in the final box.

3.3.2 Sequence Differences Between Homologues

The differences between the amino acid sequences of the MHpara proteins include the F78I polymorphism, and in the 'shorter' version, a K39E. This latter polymorphism causes the signal peptide region to differ upstream of residue 39 as predicted by ORF Finder using the DNA sequence as illustrated in Figure 3.4 sequence 184. This short sequence is the result of a frame shift early on due to a deletion, F78I is a change from a hydrophobic aromatic residue to a hydrophobic aliphatic residue. The phenylalanine residue, as expected from a larger aromatic residue, blocks the slight pore within the protein barrel to a greater extent than isoleucine and causes a change in the predicted surface of the protein (Figure 3.24). Both lysine and glutamate are polar hydrophobic residues, although lysine is a little bigger and is charged. The lack of knowledge about the function of this protein prevents us from knowing whether this would affect the proteins' interactions. The shorter protein lacks the standard signal peptide cleavage site and is therefore not likely to function in the same manner to the other paralogues if at all. There is a possibility that it has an unusual signal peptide not recognised by the prediction software.

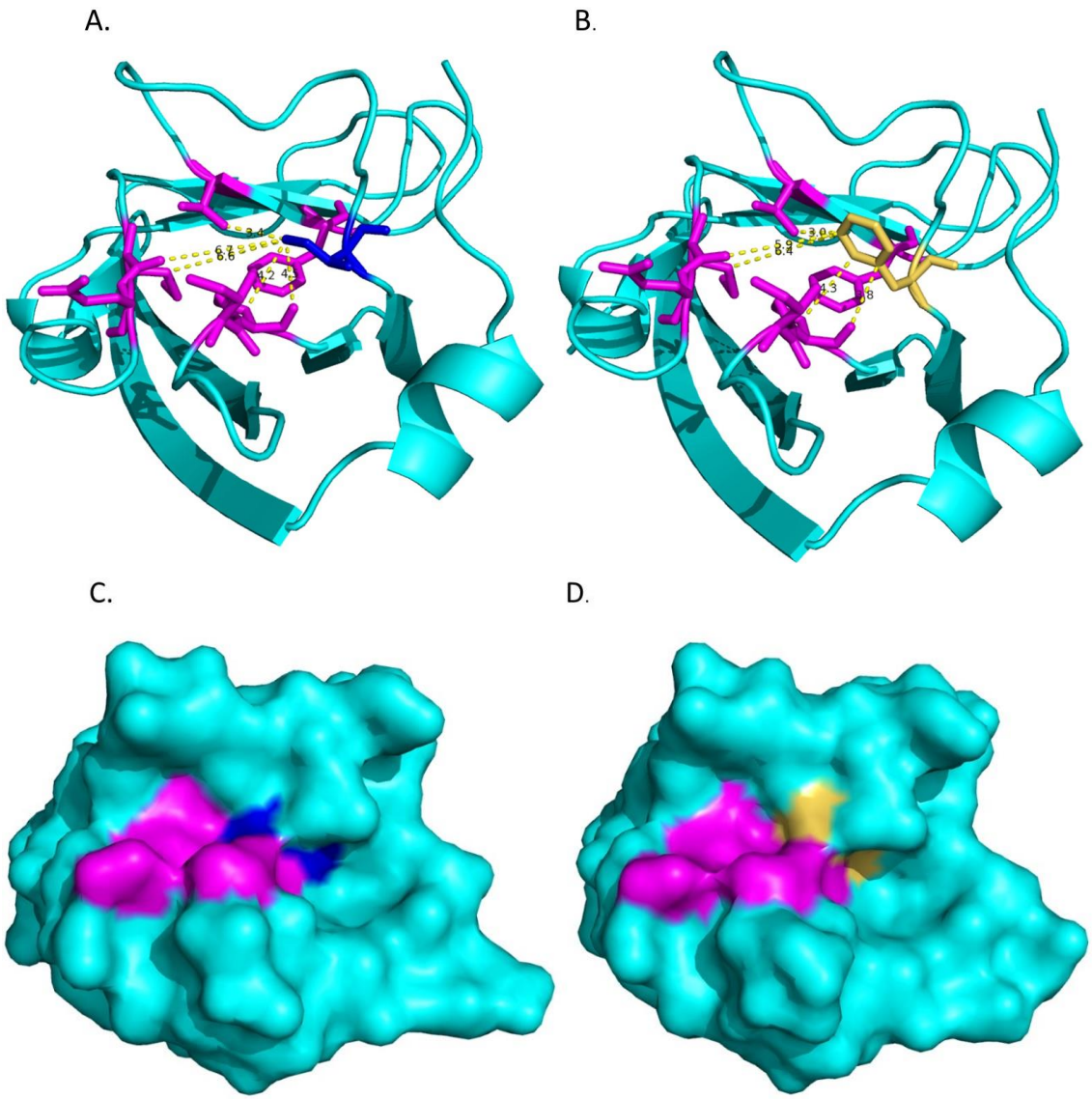


Figure 3.24 Amino Acid Substitutions in MHpara.

A. shows isoleucine in blue, and **B.** shows phenylalanine in yellow. **C.** and **D.** show the surface changes as a result of this substitution. Mutagenesis generated in Pymol and is therefore an estimation of the position of the Phe. Protein structure solved using the isoleucine variant.

3.3.3 Prediction Software

Prediction software was used as a substitute to *in vitro* assays, however, there are varying levels of accuracy of these predictions as identified by the confidence limits. In terms of signal peptide prediction, signalP was chosen based on a study (Choo et al., 2009), who found it to be the most accurate prediction tool. The predicted signal

peptides in APP and MHpara were processed as shown by N terminal sequencing, and the other proteins were assumed to be processed too, although any differences between the predicted MW to that seen on the SDS-PAGE gels could be due to the manner of protein folding.

The secondary structure prediction software, SABLE is quoted to have an accuracy of around 78% based on the protein reference sets used. Phyre2 has an accuracy of 75-80% when predicting secondary structure. This potential for error was highlighted in some of the regions in the *M. haemolytica,* where homologues with identical amino acid sequences differed from each other in the predicted secondary structure. The residues IKKAV were predicted by SABLE to be alpha helical in two of the proteins, with a beta A or AV in two of the others. This mostly agreed with the results seen with Phyre2, with a small number of differences as stated; Phyre2 predicts the region IKKAV to be helical for all proteins with none of the beta strands suggested by SABLE. This highlights the limitation of the predictive methods, and that these results should be used in conjunction with other analytical procedures.

While PSORT is reported to be the most accurate cellular localisation prediction software at 97.3% accuracy, it is still limited by the selection of reference proteins available. The MHpara results could indicate a dual location protein, but further work is required to find out how this protein functions.

3.3.4 Circular Dichroism

Circular dichroism is a technique used to estimate the secondary structure of the proteins. While software such as DichroWeb provides a good indication of the percentage of the protein that is alpha helical or beta strand, it is not 100% accurate and is again dependent on the proteins within the reference set.

This could explain why the PhuT secondary structure was estimated to be mainly beta strands, when the predicted secondary and tertiary structures suggested a majority of alpha helical structures with 6 or 8 short beta strands. Additionally, both secondary structure predictive software results showed around 46% alpha and 14% beta. This was not seen with the DichroWeb results, even when analysed with different reference sets.

The double negative peak of the spectrum suggests alpha helical structure, however, the wavelength of the peak is 217 nm rather than the 222 nm seen with the average alpha helical protein. It does have the second dip at the same location as alpha helical proteins, along with the positive peak. The higher level of disordered region/ turns/ coils could suggest that the protein had not refolded correctly, although as seen in Chapter 4, the protein still functioned in a haem binding capacity.

Although both proteins purified from the supernatant and re-solubilised fraction were observed to bind haem, further analysis could be performed on the protein purified from the supernatant that was soluble under different conditions to ensure the results have not been affected by the re-solubilisation procedure. Additionally pH, temperature and solvent can affect the structure, and although the protein was stored in the same 1 x PBS at -20°C, the temperature differed between experiments, and was considerably lower for the CD spectra, at between 18 and 20°C compared to the 37°C of the peroxidase activity assay and 20- 25°C of the hemin absorbance spectra.

3.3.5 Crystallography

While SARI C-His and MHA_1205 C-His crystals did not diffract at all, MHA_0579 C-His did diffract, however, Jason Wilson was unfortunately not able to solve the structure in its 'apo' form. The phasing proved more difficult than hoped because of the lack of sequence similarity to previously solved structures, including the neisserial Gly1s. Attempts were made to solve the structure in its 'holo' form with hemin soaked crystals but to date it has been impossible to recreate the crystals to optimise this. This may be worth attempting for the MHA_1205 and SARI as they could be involved in haem binding, which may stabilise the proteins for longer periods of time and possibly allow these crystal to diffract.

3.4 Conclusion

Potential Gly1 homologues appear to be present across a wide range of Gram-negative bacteria, and the list is growing as more and more genomes are sequenced in their entirety and added to databases. Some homologues are present in the majority of bacterial isolates within a species, where others are limited to one or two serotypes. The former could present potential novel vaccine targets in the face of increasing antibiotic resistance.

Further investigations need to be conducted to confirm the cellular location of these proteins but the predicted locations suggest that MHA_0579 may indeed have a similar role to the neisserial Gly1 proteins and potentially SARI, whereas the more similar *M. haemolytica* homologues were predicted to remain in the periplasm. It is possible that these proteins could form different parts of a nutrient uptake system, rather than all having the same function.

The MHpara protein structure was a similar antiparallel beta strand structure to the neisserial Gly1 proteins. Unlike these neisserial Gly1 proteins there is a structure in the PDB database that is believed to be similar to the MHpara protein and could provide an indication of the role of this protein as a lysozyme inhibitor. The secondary structures of all other Gly1 proteins appeared to have a large percentage of beta strands, which was in agreement with them being a family of bacterial proteins. Further experiments to obtain structures of those putative Gly1 homologues, and the PhuT haem binding protein would be interesting and could provide informative insights into their functions.

Clues have been uncovered throughout the sequence and structural analysis in this chapter providing guidance as to where to begin functional characterisation of these proteins, with haem interactions providing the most interesting starting point. It is also of interest to note that some of the flanking genes of these proposed Gly1 homologues have been implicated in antibiotic resistance such as ICE proteins and acetyltransferases, which could provide clues towards the protein function in a broad sense.

4. Haem Interactions

4.1 Introduction

Iron uptake is vital for bacterial cellular survival. In low-iron environments such as the respiratory tract, bacteria use a number of mechanisms to enhance the likelihood of accessing host iron, which is often sequestered in high affinity proteins as summarised in the introduction. Haem is a major source of iron that bacterial cells can utilise. A number of haem uptake pathways in Gram-negative bacteria have been characterised to date, however, this is by no means an exhaustive list and the discovery of new proteins involved in these pathways is ongoing.

Haem acquisition as an iron source is a known virulence factor in many bacteria (Anzaldi and Skaar, 2010), therefore identification of haem binding proteins could provide potential new targets in the struggle against antibiotic resistance. Haem uptake genes have been discovered within a 'haem transport locus' in a number of bacteria, such as *S. dysenteriae* which contains eight open reading frames (ORFs) including genes for ShuA- a haem receptor and ShuT a periplasmic haem binding protein (Wyckoff et al., 1998). The presence of such a locus could be used to identify currently uncharacterised proteins as potentially involved in haem uptake.

There is no single amino acid motif consistently associated with the binding of haem to known hemoproteins, and both the method and affinity of haem binding vary depending upon each proteins' particular functional requirements. This is an important aspect of the haem uptake system, both enabling the movement of the haem ligand between proteins and in energy-providing mechanisms. Not all contacts between haem and protein are covalent, with non-bonded contacts also providing important interaction sites to stabilise the binding, for example electrostatic bonds.

A collection of residues have been identified as important in protein-haem binding on a regular basis. As mentioned in the introduction, one use of haem in a bacterial cell is incorporation into the cytochrome C protein of *E. coli* which occurs covalently using the sulphurs from two cysteines and a histidine with the motif 'CXXCH', although the histidine can be substituted and the motif still bind haem (Allen et al., 2005). Alternative

binding sites include the haem binding pockets of the PBPs ShuT and HasA which coordinate haem with a tyrosine residue (Ho et al., 2007, Wolff et al., 2008). Further haem interactions have been shown to occur between a histidine residue and the NPDL motif in the HmuR in *P. gingivalis* (Liu et al., 2006). Haem- protein interactions have also been seen via Arg and Met residues, however, alternate motifs all incorporate cysteines, for example a CXXCK and CXXXCH.

The function(s) of the Gly1 family of proteins as a whole is currently unknown. The novel protein fold of the neisserial Gly1's does not provide any further clues, as no similar proteins have previously been characterised. A possible starting point was to analyse the genes flanking each gene of interest. In the neisserial species containing Gly1, the flanking genes code for proteins involved in haem biosynthesis (see Chapter 3, Figure 3.5). In addition to this, *N. meningitidis* has been shown to interact with both haem and components of erythrocytes in previous studies in the Sayers lab group (Sathyamurthy, 2011).

The analysis of flanking genes of potential *M. haemolytica* Gly1 homologues in Chapter 3 showed that the gene for MHpara is located downstream of genes coding for a possible haemoglobin receptor, a potential periplasmic haem binding protein and hemin ABC transporter proteins in a putative haem uptake locus (Chapter 3, Figure 3.5). This points to a possible role in haem uptake, therefore we hypothesise that Gly1 proteins are involved in haem acquisition.

In a study looking at the consequences of iron limitation on gene expression in *M. haemolytica*, genes coding for potential haemoglobin receptors, a PBP and hemin ABC transporters were upregulated, as was mh0603, a putative Gly1 homologue (Roehrig et al., 2007). While this could indicate that the Gly1 homologue is involved in haem uptake, other genes that were upregulated in the study were involved in iron uptake from a variety of alternate sources. These findings further support the results observed by (Sathyamurthy, 2011) in the hypothesis that Gly1 proteins are involved in bacterial iron uptake, with haem the likely source.

The aim of this chapter is to investigate whether or not the newly proposed Gly1 homologues are also hemoproteins.

4.2 Results

4.2.1 Hemin Agarose Binding Assay

The hemin agarose 'pull down' assay has been widely used to demonstrate the ability of proteins to bind haem (Lin et al., 2013). Initial hemin pull down assays performed in the Sayers lab group with *N. meningitidis* Gly1 showed that it was enriched after incubation with hemin agarose beads. It was therefore used as a positive control for these assays when testing proteins of unknown function. PhuT, predicted to be a periplasmic haem binding protein of *M. haemolytica,* was successfully cloned and expressed, and was therefore also assessed for haem binding ability.

The SDS-PAGE gel analysis of the pull-down assays were studied on a gel-by-gel basis to eliminate gel-related variability and to uniformly interpret protein-binding ability. ImageJ software was used to measure protein band intensity in each individual lane to elucidate the ratio of BSA: protein of interest, with the two proteins together making 100% of the lane total.

BSA was used as a non-specific control to demonstrate that the hemin agarose beads had been sufficiently washed to remove any non-specifically bound proteins as evidenced by an almost total loss of BSA in each 'pellet' lane. Results were therefore seen as change in the percentage of the lane total comprised by the protein of interest. An enrichment of this protein in the 'pellet' lane compared to that in the 'stock' lane suggests hemin binding, while no increase, or a decrease, suggests no hemin binding by the protein in question (Figure 4.1 and 4.2).

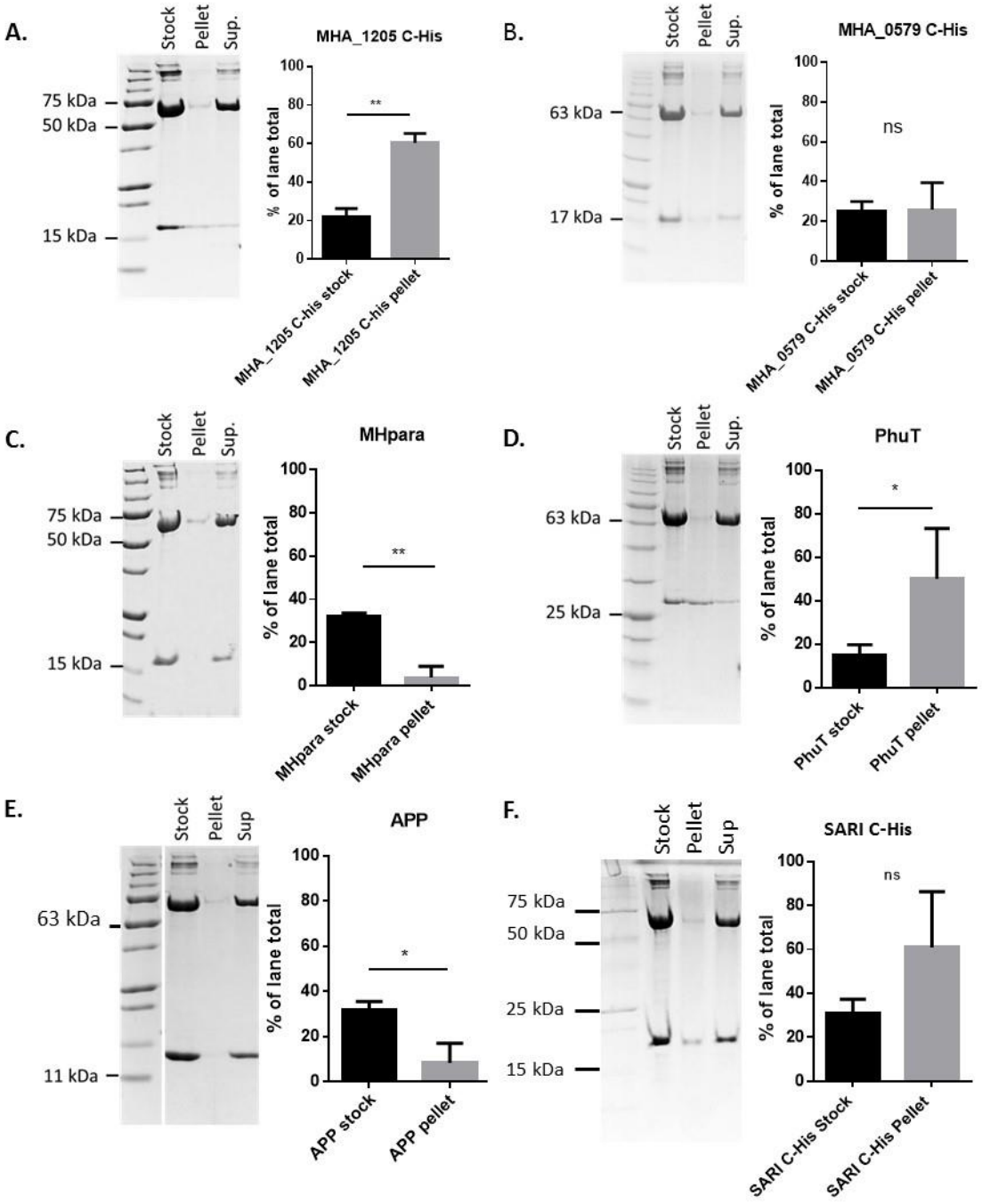


Figure 4.1. Hemin Agarose Binding Assay.

SDS-PAGE gels on the left of each panel show the pull-down, BSA is the 60 kDa band and the protein of interest is the lower MW band. The graphs show the percentage of the lane total made up by the protein of interest in the stock (black bar) and the pellet (grey bar) respectively. Of the Gly1 proteins of interest, only MHA_1205 C-His showed a significant enrichment (** indicates $P < 0.005$), with MHA_0579 C-His and SARI C-His a slight, yet non-significant, enrichment was observed. PhuT was significantly enriched (* indicates $P < 0.05$). An independent T- test was used.

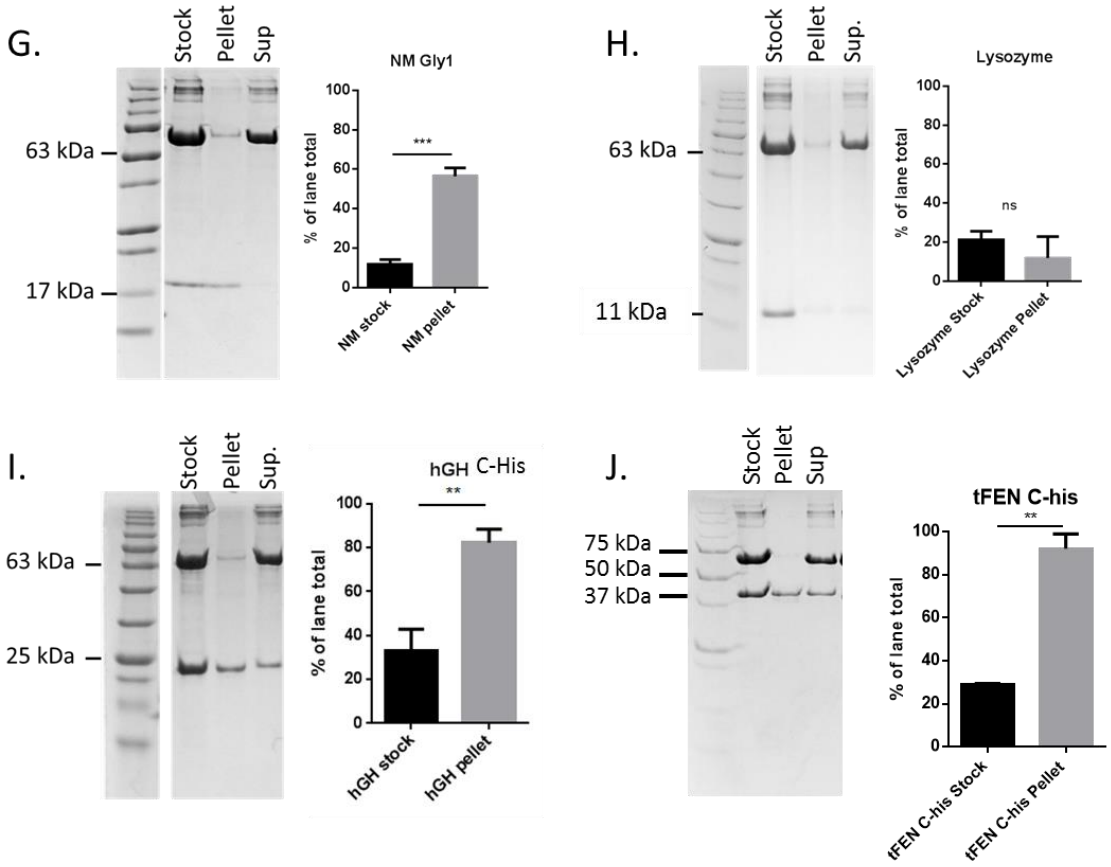


Figure 4.2 Hemin Agarose Binding Assay Controls.

Panels G-J show the controls used for this assay. *N. meningitidis* Gly1 has previously been shown to bind haem, and here was significantly enriched (*** indicates $P < 0.001$). Lysozyme was not enriched as a negative control. However, as a predicted further negative control, hGH C-His was significantly enriched (** indicates $P < 0.01$). An additional polyhistidine tagged protein, tFEN C-His, was therefore used (Panel J). This also showed unexpected significant enrichment (again, ** indicates $P < 0.01$). N = 3.

As predicted based on previous results of the Sayers group, the *N. meningitidis* Gly1 protein was significantly enriched, with the lane percentage increasing from approximately 10% Gly1 and 90% BSA in the stock lane, to approximately 55% Gly1 and 45% BSA in the bound fraction. PhuT was also found in the pellet fraction and was consistently enriched, a result shown to be statistically significant which supported the hypothesis that PhuT is a haem PBP of *M. haemolytica.* This binding also demonstrated that the protein had refolded correctly.

Lysozyme does not bind haem and was therefore used as a negative control. In this assay the percentage of lysozyme in the lane total decreased as expected for a non-haem binding protein. Contrary to the overall hypothesis for Gly1 proteins, neither the MHpara protein nor the APP Gly1 protein showed any binding to the hemin beads. MHA_0579 C-His protein showed a slight level of enrichment, however, this was not found to be significant. MHA_1205 C-His and SARI C-His both displayed a significant enrichment when incubated with the hemin beads, supporting the hypothesis that these proteins are involved in haem binding.

4.2.2 Haemoglobin Agarose Binding Assay

As the haem binding ability of proteins analysed using this assay was not conclusive, the hypothesis that the Gly1 proteins may act as hemophores was still uncertain, therefore haemoglobin agarose beads were used with the same protocol as the hemin agarose bead binding assays.

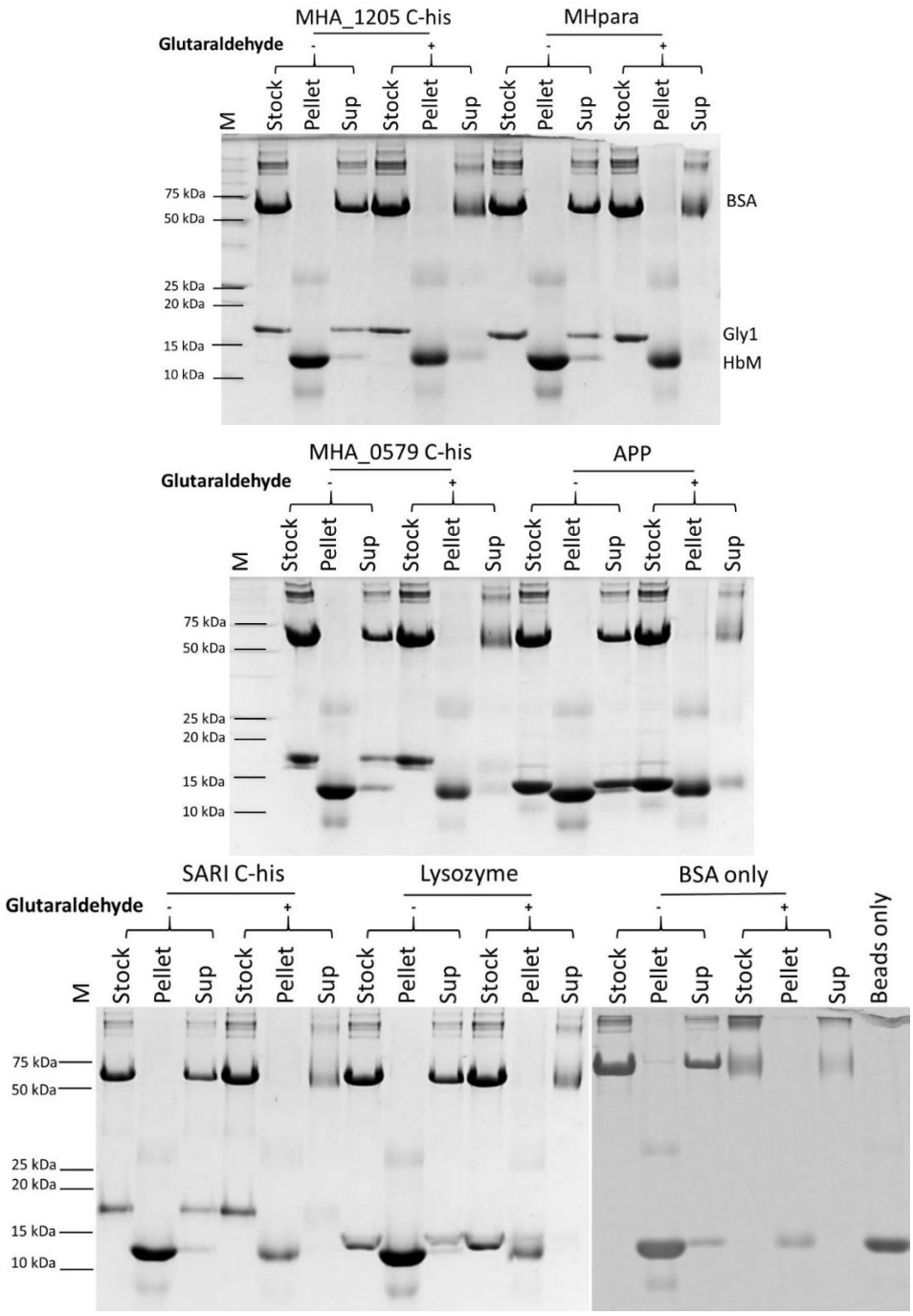


Figure 4.3. Haemoglobin Agarose Binding of Gly1 Proteins.

The panels show SDS-PAGE separation of the **stock** protein used comprised of Gly1 and BSA, which was incubated with haemoglobin agarose beads, the subsequent **pellet** after three washes and the first **sup**ernatant. Samples were incubated both with and without glutaraldehyde. The protein band at approximately 13 kDa was the haemoglobin monomer (HbM). No binding was seen. The images are representative of three repeats.

This experiment was an attempt to identify whether the proteins of interest had any haemoglobin binding ability, as they could potentially form haemoglobin receptors. Glutaraldehyde was used to immobilise any transient interactions to visualise any binding that may have otherwise been missed, as was hypothesised with the neisserial Gly1 protein. This has been observed in haemoglobin receptors to allow haem transfer from the haemoglobin and into the cell periplasm. If the proteins bound the haemoglobin even transiently we would expect to see the protein in the pellet lane along with the haemoglobin monomer. In this case, the results showed no interactions with the immobilised haemoglobin, however, no haemoglobin binding proteins were available as a positive control, as attempts to overexpress the cloned *M. haemolytica* HmbR protein in *E. coli* cells were not consistently successful (Figure 4.4).

The low MW band around 13 kDa seen in the pellet lane of each experiment was the dissociated haemoglobin monomer. Some of the haemoglobin was seen forming higher MW bands in the experiments without glutaraldehyde, but this does not suggest any interactions with the Gly1 proteins. Instead, the higher molecular weight band appears to be protein dimers based on the MW of around 26 kDa. Haemoglobin readily forms multimers, and is found in tetramer form *in vivo*. Glutaraldehyde was used at a high concentration to stabilise any transient interactions. *M. haemolytica* and *S. enterica arizonae (IIIa)* Gly1 proteins are known to form multimers (see Figure 5.5, Chapter 5), and this may have led to the Gly1 protein monomers being at too low a concentration to be observed easily in the supernatant. Only the APP protein monomer was observed in the supernatant after glutaraldehyde incubation. Densiometry was again used, however, the results showed no proteins in any of the pellet lanes other than the haemoglobin monomers.

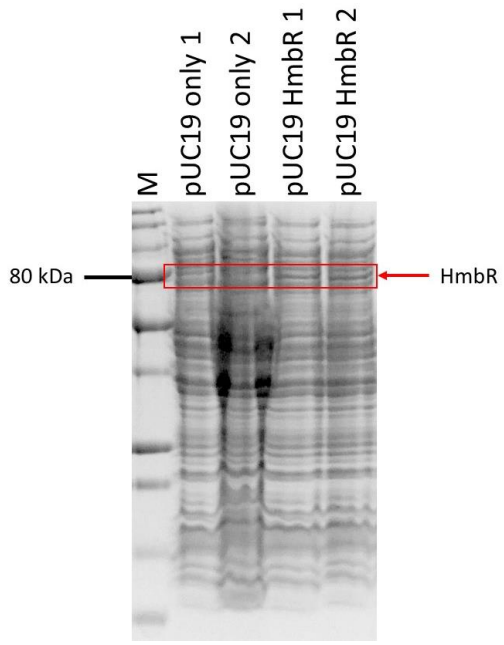


Figure 4.4 Haemoglobin Receptor Expression

The *M. haemolytica* HmbR was cloned and expressed in *E. coli* BL21 (DE3) cells. HmbR MW 79.5 kDa. The lanes show protein expression in empty pUC19 vectors and in those containing the HmbR receptor, separated by SDS-PAGE.

Expression of the HmbR protein in *E. coli* was seen in initial cloning, however, repeated attempts at expressing the protein in whole cells for Gly1 interaction analysis using techniques shown to work with the *N. meningitidis* proved unsuccessful (Figure 4.4)**.**

4.2.3 Hemin Absorbance Spectra

The hemin absorbance spectrum has been widely used as a tool to examine haem-protein interactions, and to identify new hemoproteins. The maximum absorbance (A_{max}) of hemin alone is around 385 nm, also known as the Soret Band or Peak. This maximum absorbance peak alters when hemin forms complexes with hemoproteins. The formation of bonds alters the spin of the iron atom within the hemin molecule and causes a shift in the $A_{max,}$ for example haemoglobin is characterised by its A_{max} of 420 nm and cytochrome P450 by its Soret Peak at 450 nm when in complex with haem. This is known as a red shift or a 'bathochromic shift' to a longer wavelength.

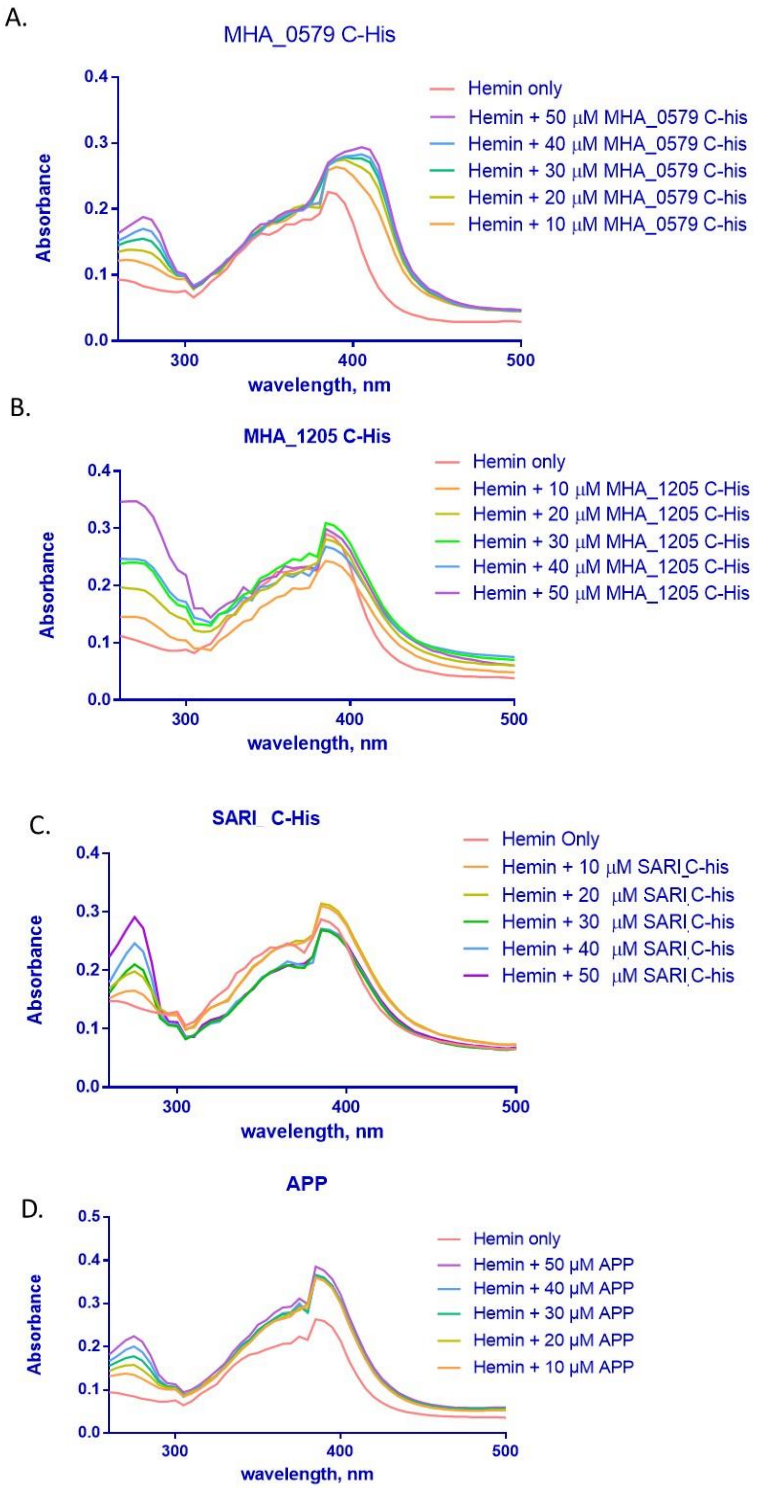
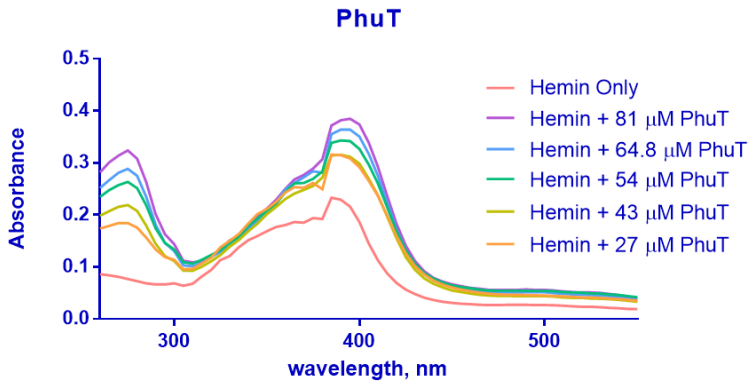


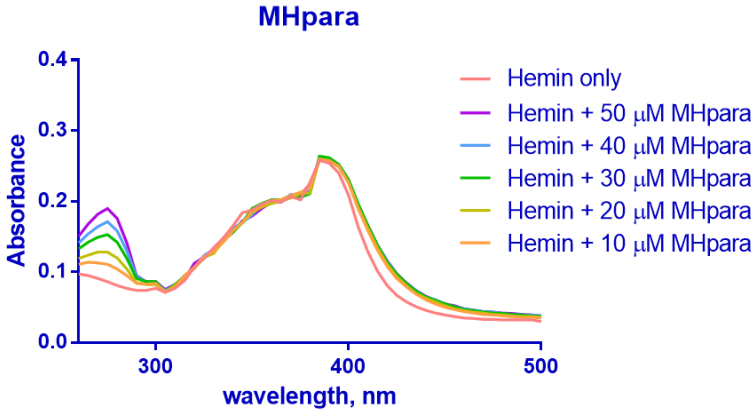
Figure 4.5. Changes to Hemin Absorbance Spectra When Incubated with Gly1 Proteins.

There was no change in the A_{max} of hemin with MHA_1205 C-His and APP, which was maintained at 385 nm with all protein concentrations. Hemin with SARI C-His had an average A_{max} at 385 nm with 10 µM protein increasing to 388.75 nm at 50 µM protein. MHA_0579 C-His increased the A_{max} from 391.25 nm to 402 nm on average with an increase in protein concentration. Graphs are representative of a minimum of 3 repeats.

E.



F.



G.

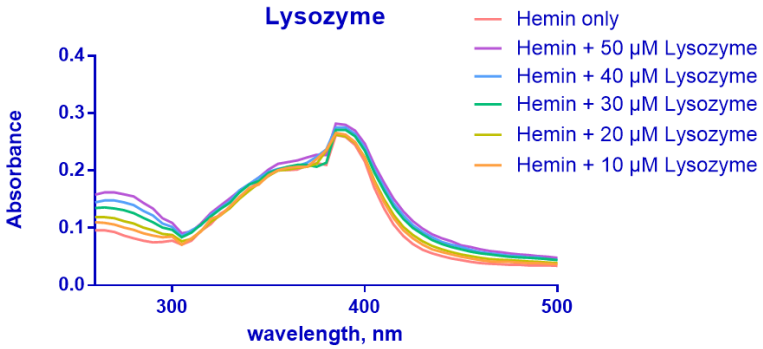


Figure 4.6. Changes to Hemin Absorbance Spectra when Incubated with Non-Gly1 Proteins of Interest.

The MHpara- hemin spectra averaged a maximum absorbance at 385 nm with 10 µM, 20 µM and 30 µM, 386.25 nm at 40 µM, and 50 µM of protein. PhuT had a maximum absorbance at 400 nm with 80 µM protein and lysozyme (negative control) had a maximum absorbance at 386.25 nm with 50 µM protein. N = 3

The results showed an increased A_{max} for hemin incubated with PhuT at the higher protein concentrations used (Figure 4.6), which is in agreement with the hemin agarose pull down data, further supporting the prediction of PhuT functioning as a hemin PBP. The MHA_0579 C-His protein, conversely to what was anticipated following the hemin pull-down assay, showed a concentration-dependent increase in the A_{max} (Figure 4.5), with concentrations up to 80 µM of protein being used (data not shown). All concentrations of protein, from 10 µM upwards, caused a clear change in the shape of the curve in addition to the change in the wavelength of the maximum absorbance at higher protein concentrations (see Figure 4.7).

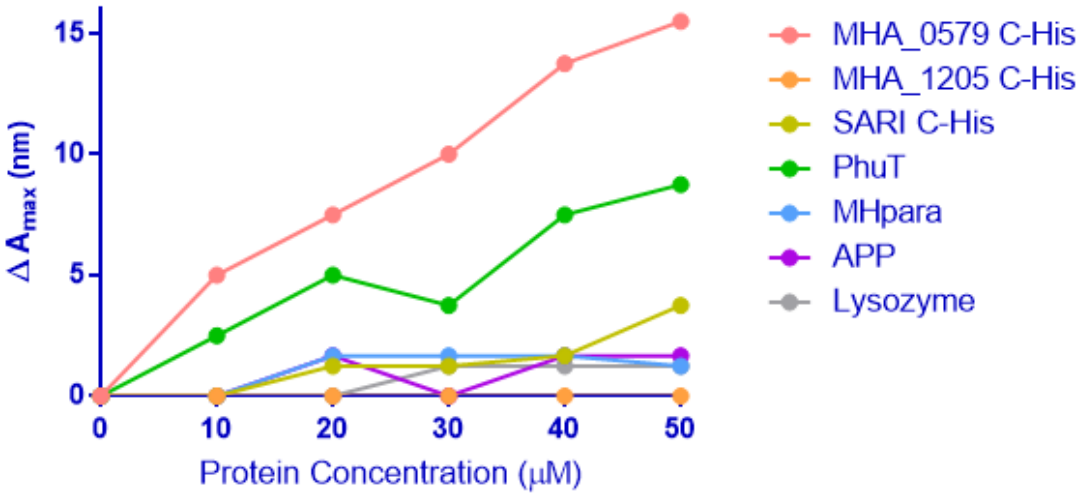


Figure 4.7 Change in A_{max} of Hemin with Increasing Protein Concentration.

The graph shows the change in A_{max} compared to hemin alone at each concentration of protein tested. Each value is an average over a minimum of 3 repeats.

As expected following the hemin agarose pull down, neither APP nor MHpara displayed any change in A_{max} and the spectra closely mirrored that of the hemin along the entire wavelength scan. This was similar to the spectra seen with the negative control, lysozyme, which does not bind haem but had an average maximum absorbance of 386.25 nm over three repeats. None of the marginal increases were found to be concentration dependent.

The spectra of hemin incubated with various concentrations of SARI C-His did not vary in shape from that of the hemin alone, however, the A_{max} increased with the protein concentration to 388.75 nm, which could indicate that some interactions were occurring to slightly alter the spin on the hemin iron. Hemin incubated with MHA_1205 C-His gave interesting results; although there was no shift in the A_{max}, the shape of the spectra differed slightly from that of hemin alone. This was most notable at the 400 - 450 nm region, where compared to lysozyme the hyperchromic shift resembled that seen with PhuT and MHA_0579 C-His. A bathochromic- hyperchromic shift in chromophores has previously been noted under differing solvent solutions, occasionally caused when the chromophore is conjugated to a ligand causing the local environment to change. This change in the shape of the curve could therefore indicate an interaction, although no binding was seen in the hemin agarose experiment after the removal of the polyhistidine tag.

Hemin is known to form complexes with itself which can affect its ability to interact with proteins e.g. the μ-oxo dimer that forms at basic pH. Haem can exist in equilibrium between the dimer and monomer state depending upon the surrounding environment, and some hemoproteins have been shown to bind both forms (Smalley et al., 2001). Hemin can be re-suspended in Triton-X or DMSO to reduce this aggregation, however, as binding was seen in the universally used PBS solution at pH 7.5 with both the PhuT positive control and the MHA_0579 C-His protein, this was not deemed necessary.

4.2.4 Haem Absorbance Spectra

The difference between haem and hemin is the oxidation state of the iron at the centre of the molecule. In hemin, the oxidation product of haem, the iron is in ferric form, i.e. Fe^{3+}, while haem contains ferrous iron, Fe^{2+}. Ferric iron is more stable than ferrous as it is already oxidised. To ensure the oxidation state of the hemin did not affect the interactions, the absorbance assays were also performed under reducing conditions using sodium dithionite and de-gassed buffers.

While the maximum absorbance of hemin alone shifted from 385 nm to 390 nm when reduced to haem, haem incubated with APP and MHpara again showed no change in the A_{max} compared to that of the haem alone. The average Soret Band with SARI C-His

present showed a hypsochromic shift, decreasing slightly to 387.5 nm (see Figure 4.8). This has been observed in the haem binding protein Irr from *Bradyrhizobium japonicum*, which works in a manner similar to Fur to regulate haem biosynthesis (Qi et al., 1999). MHA_0579 C-His again caused a bathochromic shift to 400 nm with haem, however, PhuT did not change the A_{max}. This could be attributed to the fact that similar PBPs (e.g. ShuT) interact with hemin Fe^{3+}-O via a tyrosine residue (Eakanunkul et al., 2005) rather than haem. MHA_1205 C-His incubation with haem caused a slight red shift to 392.5 nm on average, suggesting an interaction was occurring.

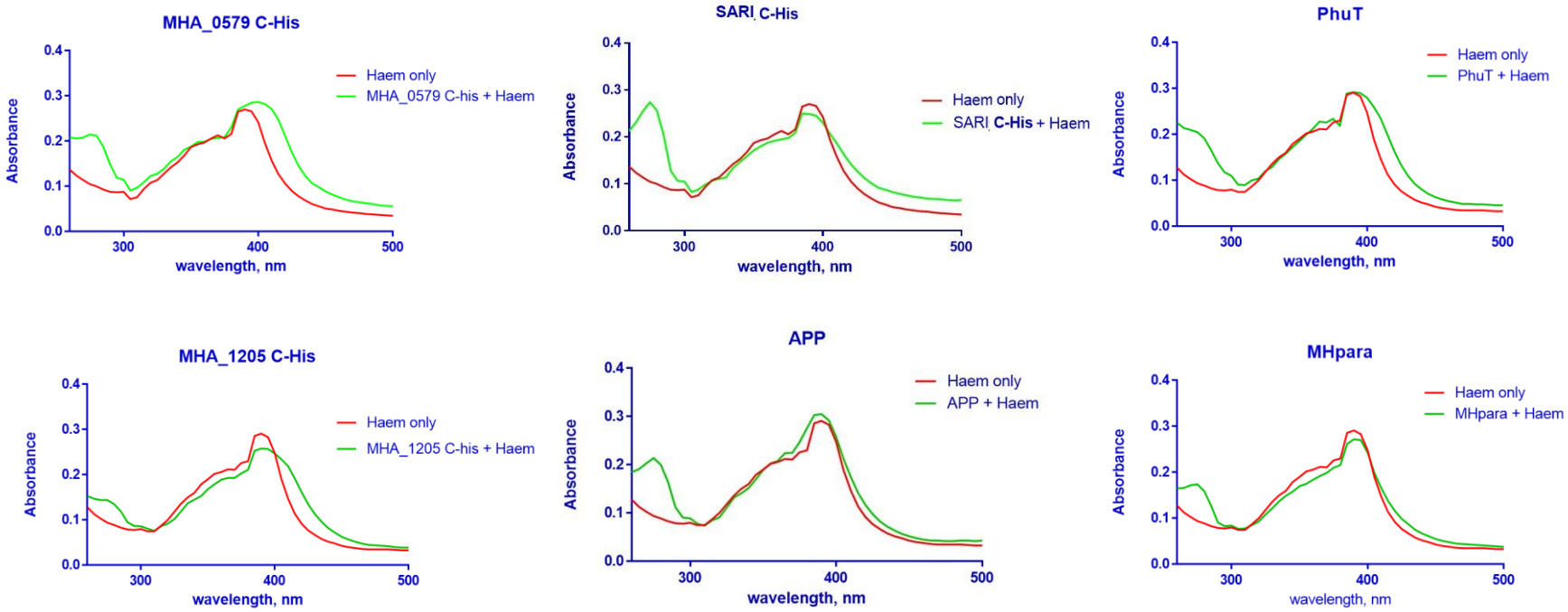


Figure 4.8. Absorbance Spectra of Haem when Incubated with Recombinant Proteins of Interest.

The A_{max} of haem in this study was at 390 nm. The A_{max} of APP and MHpara also remained at 390 nm. MHA_1205 C-His showed a slight A_{max} shift to 392.5 nm on average, SARI C-His shifted to 387.5 nm and MHA_0579 C-His to an average A_{max} of 400 nm. The A_{max} of PhuT remained a 390 nm. N = 3.

4.2.5 Hemin Peroxidase Activity

Hemin has intrinsic peroxidase activity due to its redox potential. This is maintained when bound by proteins and has been used to characterise novel hemoproteins (Lin et al., 2013, Lee et al., 2007). An ELISA-style method was used to analyse any binding to the proteins of interest. By coating the 96-well plate with the proteins of interest and adding hemin to this, any unbound hemin would be removed by the washing process, leaving only protein-bound hemin to react with the TMB substrate. Therefore an increase in absorbance compared to the hemin-only control would suggest hemin-protein interactions (Figure 4.9).

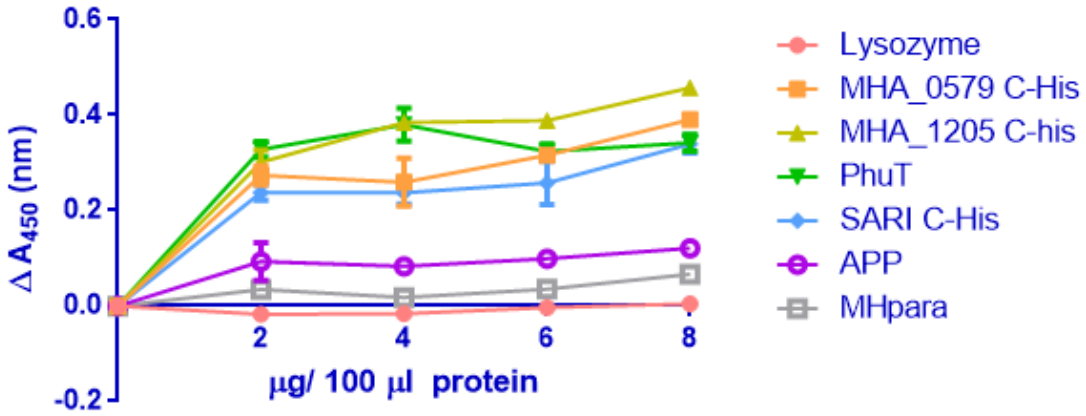


Figure 4.9. Hemin Binding as Demonstrated by Peroxidase Activity.

Peroxidase activity shown here as a change in absorbance normalised to the hemin only value. The peroxidase activity of haem increased in a concentration-dependent manner with MHA_0579 C-His, MHA_1205 C-His, SARI C-His and PhuT. Each showed a significant increase at 8 µg protein compared to lysozyme ($P<0.0001$). The negative control of lysozyme showed no activity, and MHpara and APP proteins showed low levels of peroxidase activity ($P<0.01$ and $P<0.001$ at 8 µg protein respectively). N = 3, graph is representative of repeats, each performed in triplicate. Independent T test used.

Lysozyme was again used as a negative control, showing no change in absorbance when normalised to hemin alone. MHpara and APP showed very little change in absorbance which supports the results of the previous experiments where no clear binding to haem or hemin was observed. The wells pre-incubated with PhuT showed a very significant increase in peroxidase activity after addition of hemin compared to hemin alone, in

accordance with the binding to hemin agarose beads and the shift in absorbance spectra seen in previous experiments.

Samples containing MHA_0579 C-His showed a significant concentration-dependent increase in peroxidase activity, again supporting the hemin and haem absorbance spectra which showed binding. The SARI C-His also displayed a concentration dependent increase in peroxidase activity, albeit to a lesser degree. Interestingly, MHA_1205 C-His also showed a significant increase in peroxidase activity which was not anticipated following the absorbance spectra results.

4.2.6 Hemin Agarose Binding Assay after Enterokinase Treatment

A number of amino acids are known to be involved in forming complexes with haem. Among these is histidine, which has been repeatedly shown to associate with haem via nitrogen to the iron axial ligand (Wolff et al., 2008, Izadi-Pruneyre et al., 2006). To investigate whether the polyhistidine tag on either the MHA_0759 C-His, SARI C-His or the MHA_1205 C-His was interacting with the hemin agarose beads, human growth hormone cloned with a polyhistidine tag (hGH C-His) was initially used as a negative control. Human growth hormone has not been shown to bind haem, unless broken down by pepsin and in this instance only the N terminal 1-44 fragment is able to bind haem (Spolaore et al., 2005). Interestingly, hGH C-His was significantly enriched in the pull down experiment suggesting that it was binding to the hemin agarose beads (see panel I, Figure 4.2). This could indicate that the polyhistidine tag was interacting with the hemin agarose beads in a manner that cannot be dissociated with 500 mM NaCl washes.

To further investigate this observation, a polyhistidine-tagged, truncated, human flap endonuclease (tFEN C-His) was also tested (donated by Sarah Oates of the Sayers lab group). As a well characterised DNA binding protein this has not been shown to interact with haem in a biological setting, however, in this assay the tFEN C-His protein was also significantly enriched suggesting binding was occurring.

To investigate whether it was indeed the polyhistidine tag that was interacting with the hemin agarose beads, enterokinase was then used to cleave the polyhistidine tag off the

Mannheimia and *Salmonella* Gly1 proteins. hGH and tFEN proteins lacking the polyhistidine tag were obtained from members of the Sayers Lab.

As shown in Figure 4.10, any enrichment seen with C-terminal polyhistidine-tagged versions of hGH, SARI and MHA_0579 proteins (see Figures 4.1 and 4.2) was completely abolished with the removal of the polyhistidine tag, suggesting the tag was interacting with the beads. Interestingly, the tFEN protein still appeared to bind the hemin agarose beads, showing significant enrichment comparable to that of the polyhistidine tagged version. MHA_1205 C-His showed significant enrichment yet when the histidine tag was cleaved, this enrichment did not occur. A previous member of the Sayers lab tested *N. meningitidis* Gly1 lacking the C terminal polyhistidine tag and found that this did not affect the haem binding in this experiment (Sathyamurthy, 2011).

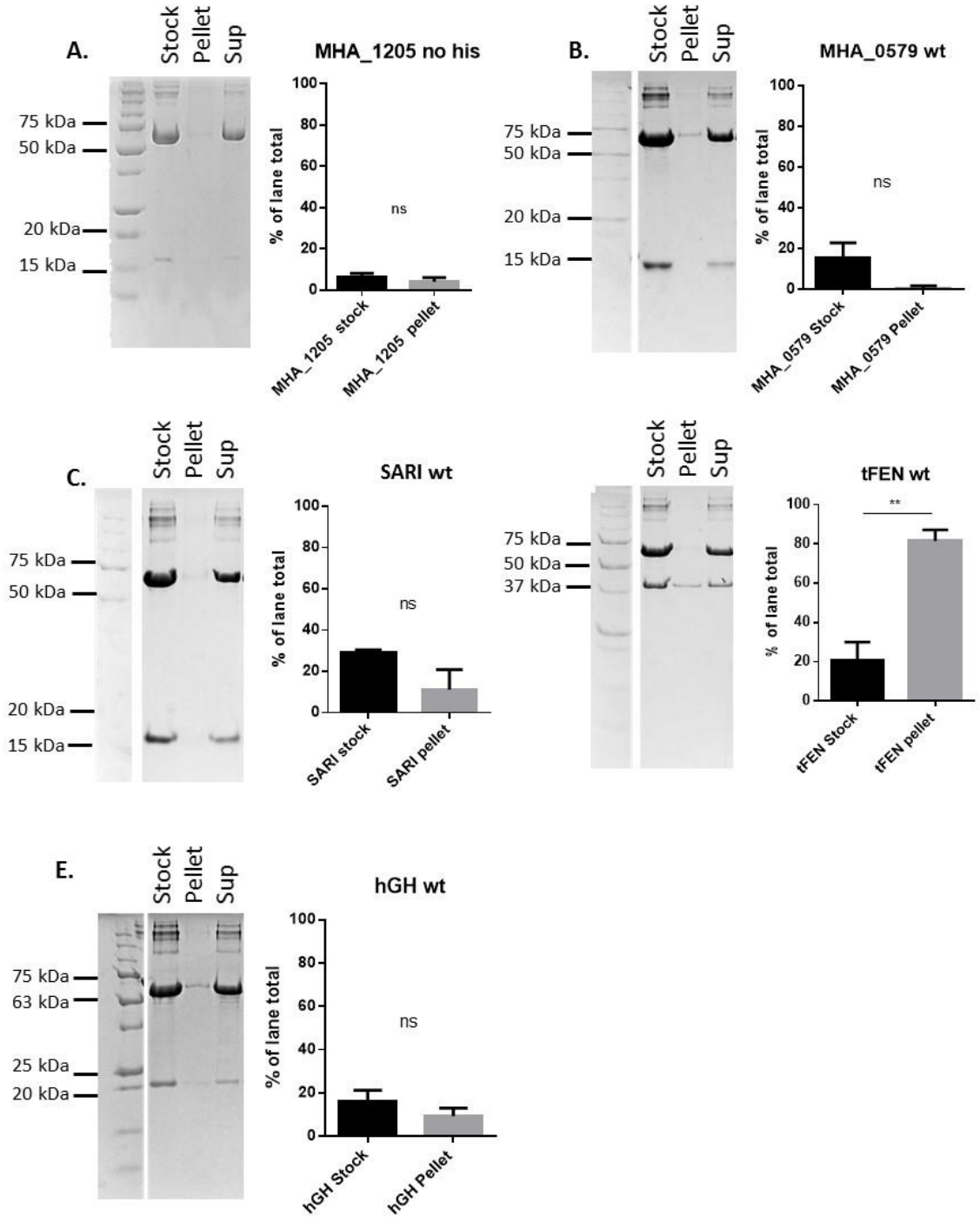


Figure 4.10 Investigating Interactions of a Polyhistidine Tag with Hemin Agarose Beads.

A repeat of the hemin agarose pull down using the same recombinant proteins either cloned with no tag or with the polyhistidine tag cleaved using enterokinase treatment. The tFEN wt protein was significantly enriched ($P < 0.005$), however, SARI wt, MHA_0579 wt, MHA_1205 wt and hGH were not. N = 3.

4.2.7 Secondary Structure of Proteins Incubated with Hemin

Circular dichroism can be used as a tool to analyse any changes in secondary structure that may arise from proteins interacting with ligands. Hemin alone does not have a CD signal, as was tested in each new experiment, therefore the only spectral effect it can generate is if its interactions alter the protein environment in some way. The results show that there was no great change in the shape of each proteins' absorbance spectrum when incubated with hemin, the only differences being in the slight increase or decrease in the amplitude of some curves.

DichroWeb was again used to estimate the percentages of alpha and beta structures along with the turns and disordered regions in each sample both alone and after incubation with hemin. Averages were made of at least 3 individual repeats for each condition and protein. The estimated percentage of MHpara secondary structure comprised of alpha and beta did not change at all between the two conditions. The percentage of alpha helices purported to comprise the PhuT protein decreased slightly from 18.9% to 13.99%. As discussed in Chapter 3 (section 3.1.3) the proportion of alpha estimated to form the majority of the PhuT structure by CD spectral analysis was not in agreement with either that reported by predictive software, or the similar amino acid sequence to other proteins with a high alpha content. The average percentage of alpha forming SARI C-His increased by approximately 5%, whereas the beta decreased from 34.67% to 31.44%. When incubated with hemin, MHA_0579 C-His saw an increase in the percentage of beta strands from 34.87% to 38.57%, whereas the alpha content decreased by 5%. The percentage of beta strands estimated to form MHA_1205 C-His increased by 5% and the alpha helix proportion decreased by 3.6% following incubation with hemin. The percentage of disorder and random coils did not change by more than 2% (Figure 4.11).

APP had shown no interactions with haem or hemin throughout the previous experiments so was not included for secondary structure analysis with hemin. MHpara was included as the structure has been solved by X-ray crystallography, and we therefore had a better idea of the accuracy of the secondary structure analysis than for

the other proteins. This did not change secondary structure at all when incubated with hemin (Figure 4.12).

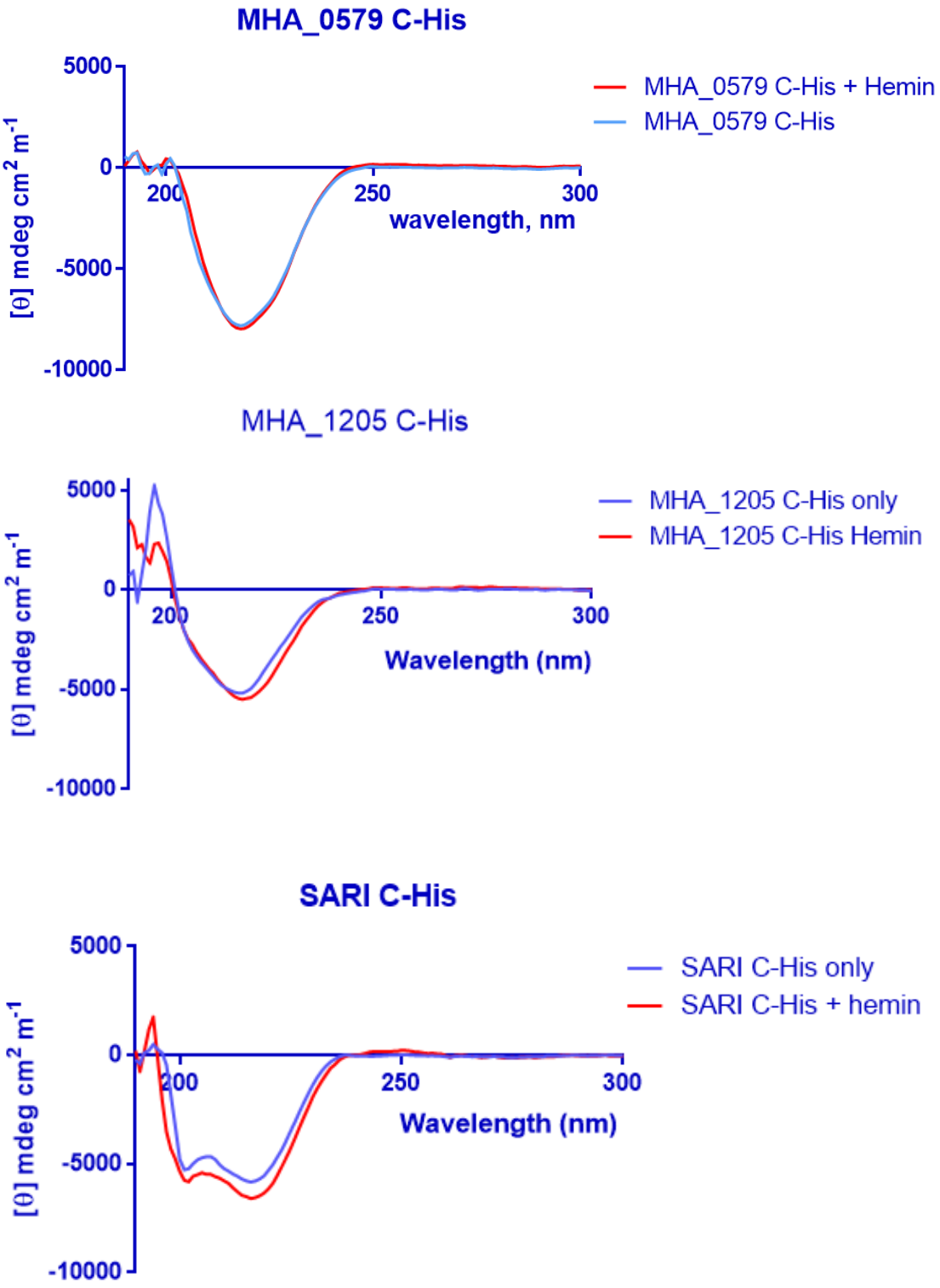


Figure 4.11 Gly1 Homologues Seondary Structure when Incubated with Hemin.

The CD spectra of Gly1 homologues was analysed with and without hemin. N = 3.

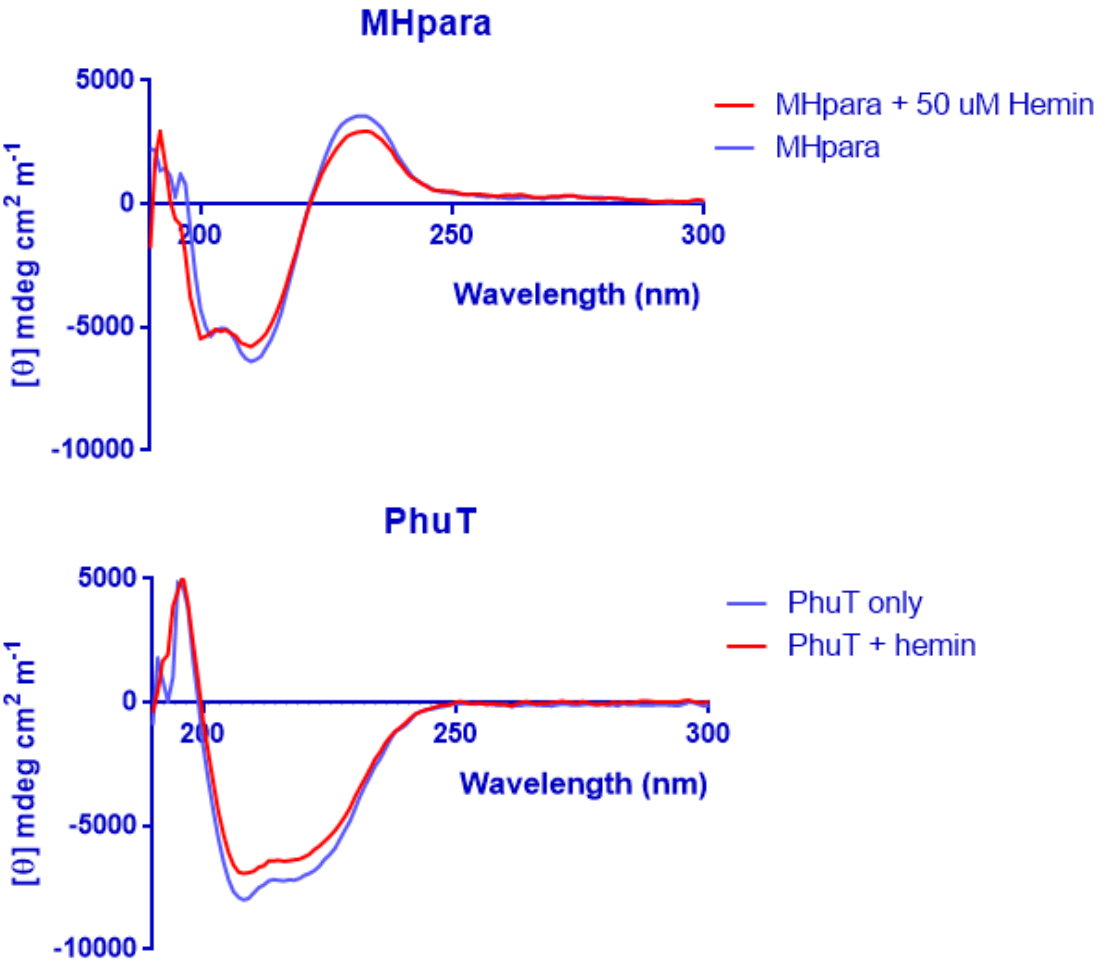


Figure 4.12. Comparison of Secondary Structure of PhuT and MHpara when Incubated with Hemin.

The CD spectra of proteins both with and without hemin. N = 3.

4.3 Discussion

4.3.1 Hemin Agarose Binding Ability of Gly1 Homologues

Immobilised hemin beads are widely used as a measure of protein-hemin interaction, but the assay is not without its limitations. The Sigma hemin agarose beads used (see Chapter 2, section 2.2.1) have the hemin immobilised via a carboxyl group (one of the propionate arms) with a spacer of 12 Å in length. A literature search revealed that the propionate arms are involved in haem binding, such as that occurring with the catalytic activity of haem oxygenase (HO) (Peng et al., 2012). With PhuT from *P. aeruginosa,* one carboxyl group of hemin is hydrogen bonded to an Arg residue (Ho et al., 2007) although the ShuT protein in the aforementioned paper did not form this bond. *P. aeruginosa* PhuT has a very similar structure to that predicted for the *M. haemolytica* PhuT which was enriched in this assay. The carboxyl group used for association to the agarose beads may therefore not always be essential for binding, but could interfere with the binding to different proteins, possibly explaining the lack of binding seen with other Gly1 homologues.

The overexpressed, purified *M. haemolytica* PhuT in its native form did not have a polyhistidine tag, so any enrichment shown with the protein in this assay supports the idea that it is indeed a *M. haemolytica* hemin PBP. The fact that no polyhistidine tag-cleaved Gly1 homologues nor the hGH protein were enriched after incubation with the hemin beads suggests that either these proteins do not interact with hemin (as expected with hGH), that the enterokinase damaged the proteins, or that the method with which the hemin is immobilised onto the agarose beads hindered any binding. The pH of the protein solution after enterokinase treatment was checked and was found to be pH 8, and therefore should not have affected the reaction.

An interesting finding was that of the non-tagged tFEN enrichment on the hemin beads. Investigations revealed the potential cause to be the presence of three histidine-rich regions in the tertiary structure (see figure 4.13), which could be associating with the hemin agarose beads, discrediting this protein as a negative control for haem binding assays, further highlighting limitations of this experiment. In addition to the histidine

rich nature, the protein also contains a number of cysteines that could act as non-specific hemin agarose ligands under theses experimental conditions.

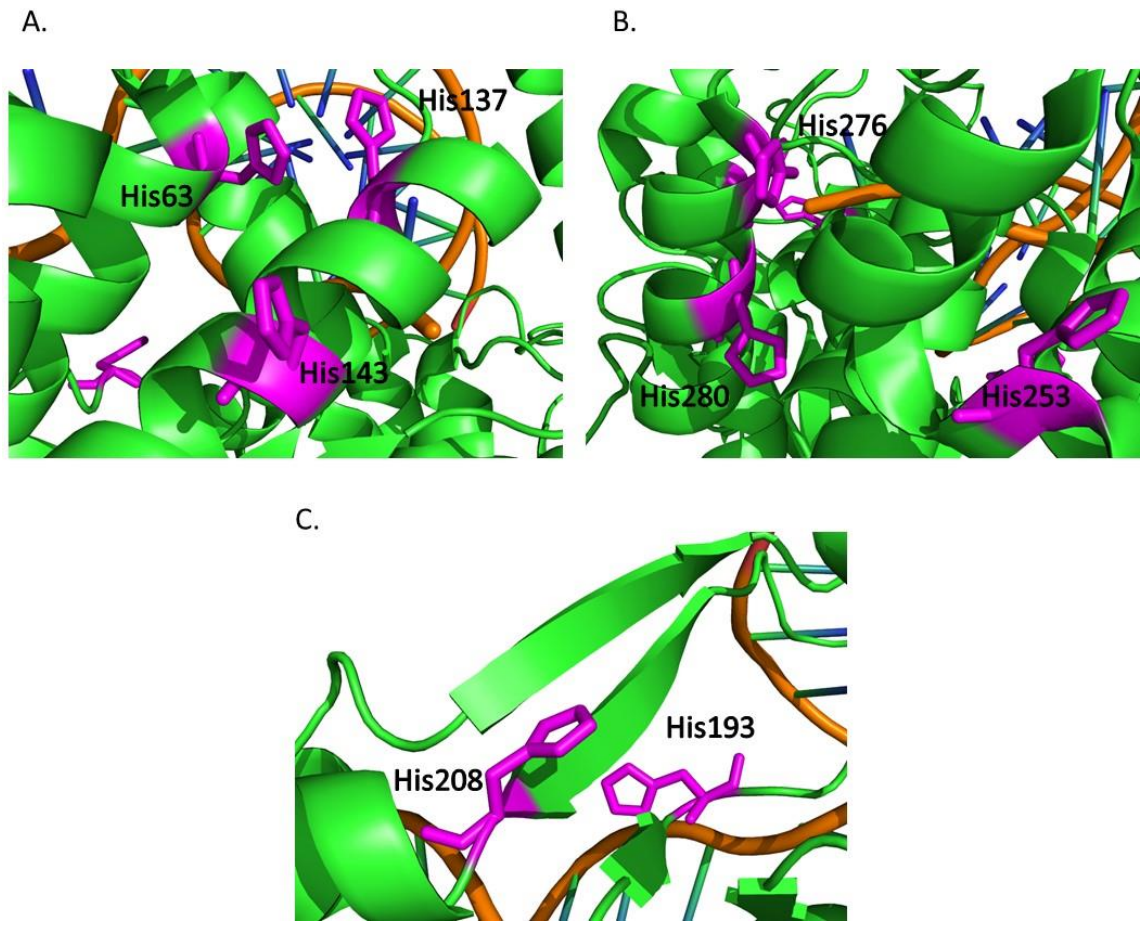


Figure 4.13 Histidine Rich Regions in FEN-1.

Panels A, B and C show the three histidine rich regions of the human FEN in complex with DNA (histidines in purple). Structure taken from RCSB PDB reference 3Q8K and image edited in PyMol.

A number of the previous studies in the literature do not use the control of a non-haem binding protein such as BSA or lysozyme to ensure any binding seen is specific, rather than weak interactions with hydrophilic regions of the hemin. In addition, during the optimisation of this protocol it was discovered that to achieve maximum reduction of non-specific binding of BSA to the hemin beads, a high concentration of NaCl in the wash buffer solution was necessary. This was shown by the almost complete removal of BSA from the pellet fractions when compared to the stock fraction.

4.3.2 Haemoglobin Agarose Binding

The limitations of the haemoglobin agarose pull down assay again included that the method by which the haemoglobin was linked to the bead matrix was via an amine bond, although more specific information was unavailable. The immobilisation of the haemoglobin could have interfered with interactions with other hemoproteins, or possibly the quaternary structure of the haemoglobin itself. After attempts to clone and express the *M. haemolytica* HmbR were unsuccessful, no positive control was available to confirm the assay was functional. The conclusions that can be drawn from this are that the proteins of interest do not bind to haemoglobin when immobilised on these particular beads.

Further haemoglobin interaction experiments were not conducted at this time due to, firstly, the predicted locations of the proteins of interest being either periplasmic or outer membrane (see Figure 3.4, Chapter 3). Secondly, while OM proteins could present plausible receptors for haemoglobin, the single unit haemoglobin receptors discovered to date are much larger than the Gly1 homologues (around 70 kDa) , and those that are smaller form bipartite receptors with other proteins. Some bipartite receptor single unit proteins can interact with haemoglobin on their own, such as HgbA in APP, however, this is again a much larger protein than the Gly1 homologues (approximately 100 kDa).

4.3.3 Hemin/ Haem Absorbance Spectra

To support results seen in the hemin agarose pull down study, further haem binding experiments were performed. Proteins alone had their absorbance spectra measured to ensure they were not contributing to each spectrum (data not shown), as some hemoproteins are purified in their holo-form. Although this would show haem binding, it would not be controlled within the confines of the experiment.

Previous work in the Sayers lab suggested that the *N. meningitidis* Gly1 bound hemin in a 2:1 ratio of Gly1: hemin (Sathyamurthy, 2011), therefore this ratio was used as the basis of this assay. Many hemoproteins bind haem in a 1: 1 ratio, however, this is not universal and, as previously mentioned, some proteins are also able to bind haem dimers. HasA has been suggested to bind in a 1:1 ratio but is secreted as both a

monomer and dimer, with the dimer domain swapping to complex the hemin (Czjzek et al., 2007). Therefore the concentration of protein used in the studies was varied from 1: 2.5 ratio of protein: hemin, to 2: 1 to encompass these potential variations in binding stoichiometry.

A decrease in the peak absorbance between 350 and 450 nm seen in hemin incubated with APP and PhuT (Figures 4.5 Panel D and 4.6 Panel E, respectively) could be attributed to an increase in aggregation of hemin over time (Srinivas and Rao, 1990). However, all spectra were measured within minutes of each other when incubated with each protein, and the only difference that may therefore have affected the results is the lower overall absorbance at these wavelengths when compared to hemin alone. This should not have affected the wavelength at which the A_{max} occurred, just the value of the A_{max} itself. An overall decrease of the absorbance spectra could indicate pipetting error when adding the hemin, however, this was not seen in the replicates performed.

The sharp dip in absorbance observed in all spectra including that of hemin and haem alone, between 380 and 382 nm, can be attributed to an artefact of the spectrophotometer at the point of transition from visible to UV light. The spectrum was found to be the same in both quartz and plastic cuvettes, and the dip did not interfere with the A_{max} in any of the spectra as checked by reducing the wavelength step to 0.5 nm from 5 nm (data not shown). The A_{max} in these experiments was 385 nm for hemin only- a higher wavelength than the dip.

The slight shift in the spectral shape of hemin and MHA_1205 C-His could be showing some non-covalent interactions that do not involve the iron of the hemin, such as the hydrophobic interactions that are often present to stabilise the covalently bound hemin into the haem binding pocket of proteins such as haemoglobin. The decrease in the A_{max} of hemin with SARI C-His could be indicative of a hexa-coordinate binding rather than penta-coordinate interaction, possibly exhibiting this unexpected spectral change. The discrepancies between the results of the two hemin binding assays discussed so far could be explained by the limitations of each experimental procedure, so a further experiment was conducted to add to the previous results.

4.3.4 Hemin Peroxidase Activity Assay

The increase in peroxidase activity of PhuT and MHA_0579 C-His when incubated with hemin support binding data from the hemin absorbance assays. PhuT has shown positive results indicating hemin binding with all three experimental techniques. MHA-0579 C-His, however, displayed results consistent with hemin binding in two of the three, showing no binding to the hemin agarose beads. MHA_1205 C-His, which only showed binding before removal of the polyhistidine tag in the pull down assay and minor changes in the absorbance spectra of haem and hemin, showed an increase in peroxidase activity. This indicated haem binding as neither the protein alone nor hemin alone showed peroxidase activity. As binding was seen in this assay, conducted in the same PBS buffer as the hemin absorbance assay, we can conclude that aggregation of the hemin, if present, did not interfere with interactions and is unlikely to be causing the different results observed with the absorbance spectra. The lack of peroxidase activity of APP and MHpara proteins supports the previous binding data, and together they suggest that these proteins are not involved in bacterial haem/ hemin uptake.

The concentration-dependent increase in peroxidase activity seen with SARI C-His is at odds with the lack of binding seen in the pull down assay post polyhistidine tag-cleavage, however, there was some enrichment with the tag still present. In addition, there was a slight increase in the A_{max} of the hemin absorbance spectra with SARI C-His, which could suggest some interaction. The maximum ratio of protein: hemin used in the peroxidase activity assay was 1: 2.5, which is consistent with the lowest concentration of protein used in the absorbance spectra assays, therefore this difference in binding results cannot be attributed to an excess of hemin in the peroxidase activity assay. A possible contributing factor is that in the hemin peroxidase assay, the higher temperature of 37°C is favourable to the haem binding reaction. The spectrophotometry had to be conducted at room temperature which fluctuated between 18 and 21°C.

The enterokinase used to cleave the polyhistidine tag off the proteins specifically recognises the DDDDK sequence, cleaving after the lysine residue. This should not have targeted any other section of the protein and should therefore not have affected the haem binding ability of SARI to the agarose beads, however, it is still a possibility worth

some consideration as a theory explaining the differences in binding seen between assays.

4.3.5 Secondary Structure Changes in Hemoproteins

We know that the tertiary structure of proteins changes upon haem binding, as seen in HasA, where the loops between beta sheets move to complex the haem molecule (Izadi-Pruneyre et al., 2006). The secondary structure of some proteins has been shown to vary when they are in the holo form or ligand free state, such as cytochrome b562 which displays a higher proportion of disorder in helix IV in the apo form compared to the holo form (D'Amelio et al., 2002). The data shows that if there is any hemin binding to the proteins of interest, the interactions do not cause dramatic changes in the secondary structure.

4.4 Conclusion

These haem binding studies have provided evidence for the functional activity of PhuT which had not yet been characterised. This supports the hypothesis that based on amino acid sequence, secondary structure estimation and predicted tertiary structure this is a *Mannheimia haemolytica* periplasmic haem binding protein. Further studies to investigate this mechanism could look at the binding affinity of PhuT for haem and subsequent interactions with ABC transporter proteins, as the genes coding for both this and the likely HmbR protein are predicted to reside in the same genetic locus.

The haem binding data presented in this chapter overall (See summary Table 4.1) provide evidence to support the hypothesis that MHA_0579, MHA_1205 and SARI interact with haem or hemin in some way, and are therefore involved in bacterial haem uptake. Despite the indications provided by the flanking genes that it may be situated in a putative haem uptake locus in the genome, MHpara did not display any haem binding activity. In addition, no haem binding activity of the APP protein has been observed, suggesting it has a different function.

	Experiment				
Protein	**Pull Down**	**Absorbance Assay**		**Peroxidase Activity**	**Secondary structure changes?**
		Hemin	**Haem**		
MHA_0579	N	Y	Y	Y	Y
MHA_1205	N	N	Y	Y	Y
MHpara	N	N	N	N	N
APP	N	N	N	N	N
SARI	N	Y	N	Y	Y
PhuT	Y	Y	Y	Y	Y

Table 4.1 Summary of Haem Binding Data.

Hemin pull down data taken from proteins post polyhistidine tag removal.

Steric hindrance may have played a role in the discrepancies between hemin agarose interactions and those seen in the other assays. It cannot be ruled out that the polyhistidine tag had an effect on the absorbance and peroxidase binding studies, however, insufficient materials were available to conduct the absorbance assays with non-tagged proteins. Repeating these with native proteins would be valid further work to ensure the binding seen was not an artefact.

While the *N. meningitidis* Gly1 and SARI C-His proteins were observed in the culture supernatant after overexpression in *E. coli* cells, which initially hinted at it being secreted and therefore a possible bacterial hemophore, the *M. haemolytica* Gly1 proteins were not secreted from *E. coli* cells. This secretion was not observed in wild type *N. meningitidis* culture when the supernatant was analysed by western blot, although the protein was also not shown to be expressed in the cells under the growth conditions used. (Sayers lab, unpublished). The protein was secreted from *N. gonorrhoeae* and was also found in membrane fractions (Arvidson et al., 1999). Further to this, no *Salmonella enterica arizonae (IIIa)* was available to analyse wild type Gly1 expression so conclusions regarding secretion could not be drawn.

As *E. coli* is not the natural host for the putative *M. haemolytica* Gly1 proteins, the isolates donated by the AHVLA and the RVC were analysed for secreted proteins. The protein was not observed in the supernatant, again using western blot, in any of the strains that had been shown to carry the genes for one or more Gly1 (data not shown). However, the western blots were not able to detect the proteins in all of the isolates that had Gly1 genes detected in PCR screening (Section 6.2.3, Chapter 6). The *M. haemolytica* strains were grown in iron replete conditions, and the transcription of genes coding for Gly1 proteins could therefore have been repressed due to lack of cellular requirement for iron acquisition.

Most haem binding proteins are made up of a high percentage of α- helical structure (Li et al., 2011), however, as this is not seen in the predicted structures of the Gly1 proteins, it is worth considering other possible functions, as they could still present valid antimicrobial targets as well as those that look to bind haem based on this data.

5. Protein Interactions

5.1 Introduction

Haem uptake pathways are comprised of proteins of different function to bring scavenged haem to the bacterial cell, transport it across the outer membrane, through the periplasm and the cytoplasmic membrane and into the cytoplasm, where it is utilised or stored. While many consist only of a single component, some outer membrane receptors involved in iron acquisition are bipartite, such as the transferrin receptor which is made of two separate proteins (Yang et al., 2011) or HpuAB, a bipartite haemoglobin receptor analogous to the TbpAB receptor.

PhuT is a suspected periplasmic hemin binding protein of *M. haemolytica*, a key component in bacterial haem uptake (Eakanunkul et al., 2005). The mechanisms of action in terms of haem uptake in *M. haemolytica* have not yet been experimentally demonstrated. Proteins of similar predicted structure and function bind haem in a 1:1 or 1:2 protein: haem ratio. With regards to protein oligomerisation, the haem PBP from *S. dysenteriae* has been crystallised as a dimer whereas the PBP for *P. aeruginosa* was found to exist as a hexameric homooligomer. The crystallisation process can, however, result in the artificial formation of dimers, so care needs to be taken when interpreting these data.

Gene duplication events can lead to protein paralogues that function in a different manner to the original protein. The MHpara protein did not appear to be functionally related to the putative Gly1 homologues based on haem interaction studies, which could be explained by divergence from the ancestral gene. The predicted function of this protein was annotated as an adhesin on the NCBI database (see Table 3.2) with the NCBI Prokaryotic Genome Annotation Pipeline (released 2013). This was based on homology with similar genes/ proteins already within the database, but as this method relies on other annotations being correct it is therefore important to experimentally investigate the potential of the protein acting as an adhesin.

Adhesins are considered one of the major virulence factors of *M. haemolytica* along with leukotoxin, LPS, capsule and the ability to acquire nutrients from the host. As

mentioned, *M. haemolytica* is surrounded by a polysaccharide capsule to protect itself from the host immune system. This capsule can mask some proteins leading to differing levels of OM protein exposure, which can affect their recognition by the host. Adhesins of *M. haemolytica* identified to date include the proteins OmpA, and Lpp1, which have been found to bind to bovine epithelial cells of the upper respiratory tract (Kisiela and Czuprynski, 2009).

To adhere to respiratory epithelial cells, *M haemolytica* are known to use a variety of mechanisms, including fimbriae and OM proteins which interact with the extracellular matrix of the host cell, such as bacterial OmpA and host fibronectin (Lo and Sorensen, 2007). The capsule may also act as a non-specific adhesin as seen in *A. pleuropneumoniae*.

Bacteria can also remain within the mucous layer or degrade the layer to improve adhesion using specific enzymes, for example in *M. haemolytica* neuraminidase degrades components of the mucous layer (Straus et al., 1993). The polysaccharide capsule of *M. haemolytica* is differentially expressed in log and stationary phases, and in *A. pleuropneumoniae* down-regulation of the genes coding for the capsule was observed in cells in planktonic growth compared to when adhered to epithelial cells (Auger et al., 2009). This reduction in capsule has been suggested as a method of improved host colonisation, although it could also expose the bacterial cell to attack by the host defence mechanisms.

Vaccines for BRD have a wide scope of potential infections to contend with. As each outbreak can be caused by multiple factors, vaccines for young feedlot animals would ideally cater to the major components involved in bacterial and viral causes of disease. Antibodies to proteins that are conserved between species would make these proteins prospective targets.

In this chapter the results of investigations into the interactions of the putative Gly1 proteins with molecules, proteins, or cells other than those related to haem binding and uptake are discussed. The previous chapter touched on Gly1-haem interactions showing possible roles for MHA_0579, MHA_1205 and SARI in bacterial haem uptake systems. Many proteins involved in haem binding function as homodimers or homomultimers for

example the tetramer haemoglobin. It is possible that these Gly1 proteins work together to form a haem uptake pathway.

5.2 Results

5.2.1 Cross Reactivity

Due to the variable level of amino acid sequence identity between the Gly1 proteins and to investigate their potential as therapeutic targets, the ability of protein-specific antibodies to recognise other Gly1 proteins was assessed. Cross reactivity of the *M. haemolytica* Gly1 antibodies was analysed initially with the MHA_1205 C-his rat antisera, the SARI C-his rat antisera and APP mouse antisera prior to acquisition of further protein-specific rabbit antisera. Anti-polyhistidine antibodies were used to both detect the polyhistidine tagged proteins for which no specific antisera were available, and also as a positive control to show the protein concentration was sufficient to elicit a response. Gly1 proteins from the four bacterial species were assessed for cross-reactivity to the aforementioned antisera. BRD is a multifactorial disease caused in part by multiple bacterial strains containing genes for potential Gly1 homologues. Thus the cross reactivity of antisera could be of huge interest in terms of drug development.

Cross-reactivity was seen between the MHA_1205 C-His antisera and one other protein homologue- *N. meningitidis* Gly1. Further cross reactivity studies using each available antisera and each available protein showed that the APP antisera, MHA_0579 C-His and MHpara antibodies did not cross react with any of the Gly1 proteins other than their respective targets, and in turn none of the other antisera reacted with the MHpara or APP proteins (Figure 5.1).

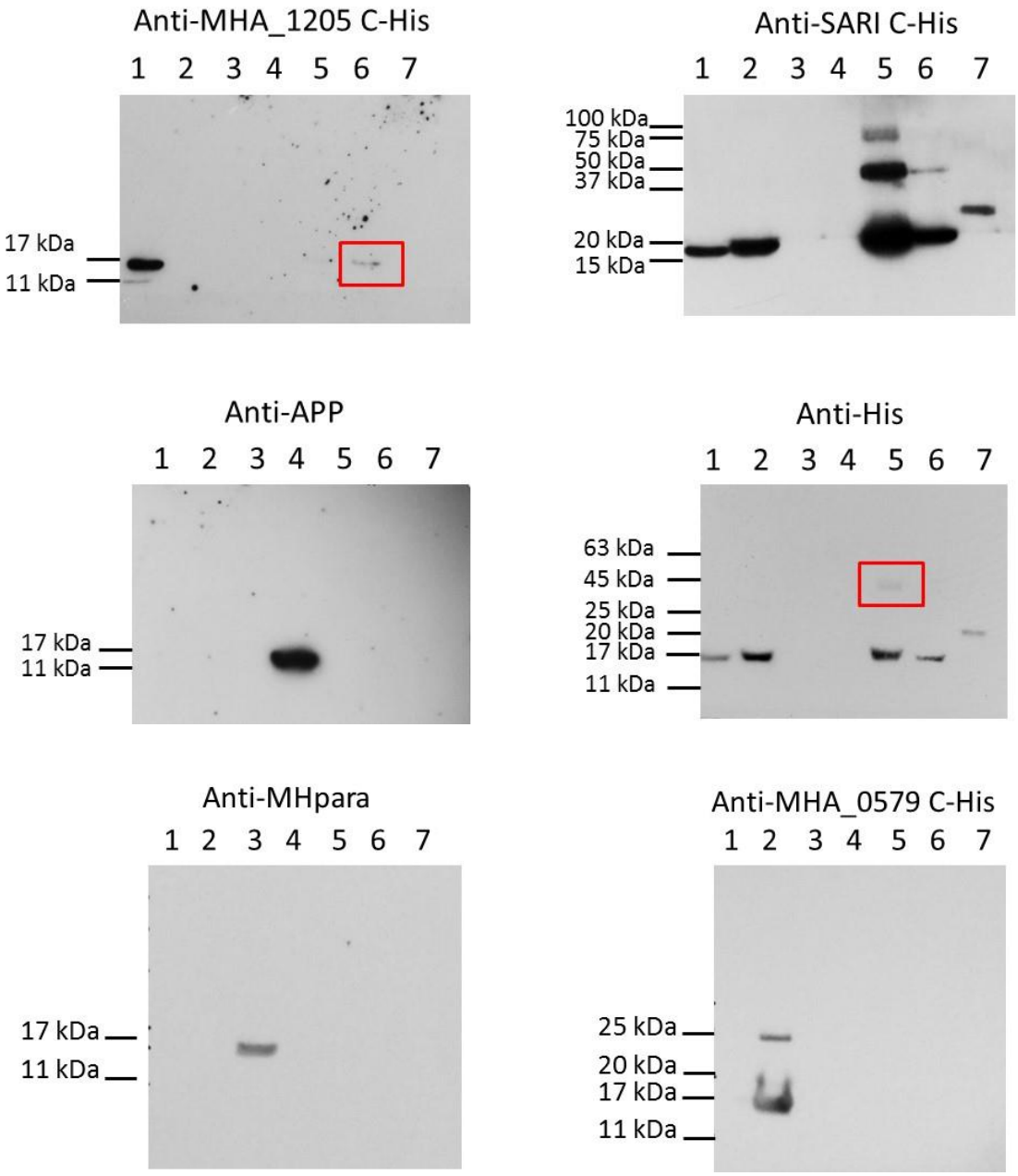


Figure 5.1. Cross Reactivity of Gly1 Antisera.

Antisera or antibodies used are stated above each image. Lanes as follows: 1 = MHA_1205 C-his, 2 = MHA_0579 C-his, 3 = MHpara, 4 = APP, 5 = SARI C-his, 6 = NM C-his, 7 = hGH C-his. Here hGH C-His was used as a negative control. In each sample, 1 mg of protein was used, diluted 1:2 with loading dye and separated by SDS-PAGE. Higher molecular weight bands observed were protein multimers, as confirmed by anti-His antibodies reacting with SARI-C-His high MW band in the red box. N=3, representative images used.

Interestingly, when exposed for a longer period of time or when using an increased concentration of protein, cross reactivity was seen using both the MHA_1205 C-His and the SARI C-His antisera to one or more other polyhistidine tagged proteins. The SARI C-His antisera reacted a lot more with the proteins possessing a polyhistidine tag than the MHA_1205 C-His antisera which only recognised *N. meningitidis* polyhistidine tagged Gly1 (Figure 5.1). The antibodies present in these antisera stocks had not been quantified therefore volumetric dilutions were made of each. In addition, neither of these antisera were affinity purified, so it is possible that the proteins were being recognised in a non-specific manner. The anti-SARI C-His antisera also reacted with the negative control of human growth hormone (hGH), suggesting that the polyhistidine tag was being recognised to some extent as the non-tagged proteins were not recognised.

In order to eliminate the polyhistidine tag as a cause of non-protein specific antibody binding in the anti-SARI C-His antisera, enterokinase was used to cleave the polyhistidine tag off said proteins. Cleavage was confirmed by SDS-PAGE separation, showing SARI C-His with a MW of approximately 17 kDa and SARI wt a MW of approximately 16 kDa (See Appendix). This resulted in a decrease of the positive results observed in Figure 5.1 in terms of recognition of the *N. meningitidis* Gly1 and MHA_0579 proteins. MHA_1205 was detected to a greater extent than MHA_0579, and *N. meningitidis* Gly1 detection was low to non- existent across repeats. No further hGH lacking the polyhistidine tag was available to assess the cross reactivity in a non- specific capacity in this assay. The SARI wt protein and the SARI C-His displayed equivalent strong reactions with the antisera. This showed that the enterokinase did not damage the protein in terms of antibody recognition (Figure 5.2).

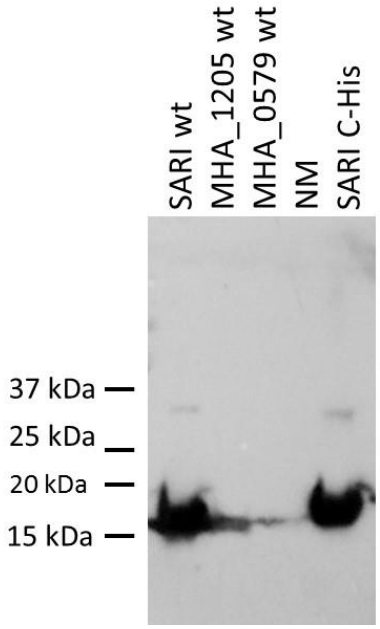


Figure 5.2. Cross Reactivity of Gly1 Antisera After Protein Enterokinase Treatment.

Proteins with polyhistidine tags removed detected using anti-SARI C-His antiserum. NM is *N. meningitidis* Gly1 wt. Protein concentrations used were 2 mg/ml for SARI wt and SARI C-His, 1 mg/ml MHA_0579 wt and MHA_1205 wt and 0.9 mg/ml NM wt. Proteins were diluted 1:2 with loading dye and separated by SDS-PAGE. N = 3.

5.2.2 Dimerisation Properties

Protein homo-dimerisation and multimer formation can occur between two or more monomer units of the same protein. The associations can drastically affect the function of the protein, for example haemoglobin comprises 4 monomer units to make a functional structure. As seen in the antibody cross reactivity assay, additional protein bands at higher molecular weights with the SARI C-His and MHA_0579 C-His proteins could indicate dimerization. The proteins were separated on denaturing protein gels, using SDS and heat uniformly, and DTT to reduce disulphide bonds for comparison.

N. gonorrhoea Gly1 was crystallised as a dimer (as shown in Figure 3.2) and recent work in the Sayers lab confirmed *N. meningitidis* Gly1 dimers and higher order oligomers in solution (Weirzbicka, 2014). *N. meningitidis* Gly1 was shown to bind haem in a 2:1 ratio (Sathyamurthy, 2011), suggesting its function is related to this dimerisation. The structure was solved by the Artymiuk and Sayers labs (Jason Wilson and Magda

Weirzbicka (Weirzbicka, 2014)) and had an almost identical structure to the *N. gonorrhoea* Gly1 protein (Arvidson et al., 2003), even crystallising with the same dimer interface. Purified samples of each *M. haemolytica* Gly1 protein were therefore assessed for homo-dimerisation properties to further investigate any similarity to the original Gly1 proteins.

Dimers can be formed by intermolecular disulphide bonds connecting cysteine residues on individual monomers. For SDS-PAGE separation, protein dimers stabilised in this manner can be disrupted using dithiothreitol (DTT) to enable greater accuracy in interpretation of purity and size. The initial dimerisation experiment consisted of separating the protein on an SDS-PAGE gel after being boiled in SDS-PAGE loading dye containing either fresh DTT or no DTT. Omitting this DTT revealed the formation of possible multimers in a number of the pure protein stocks at neutral pH (Figure 5.3).

MHA_0579 C-His formed what appeared to be dimers of approximately 32 kDa in size. The protein required a much higher concentration per sample to give visible dimers on the Coomassie stained gel compared to MHA_1205 C-His and PhuT. SARI C-His also showed some higher molecular weight bands, potentially dimers. PhuT, with a monomeric molecular weight of 30.7 kDa, formed visible multimers at low protein concentrations on the SDS-PAGE gel seen as a band at approximately 60 kDa. No multimers were observed with APP or MHpara proteins.

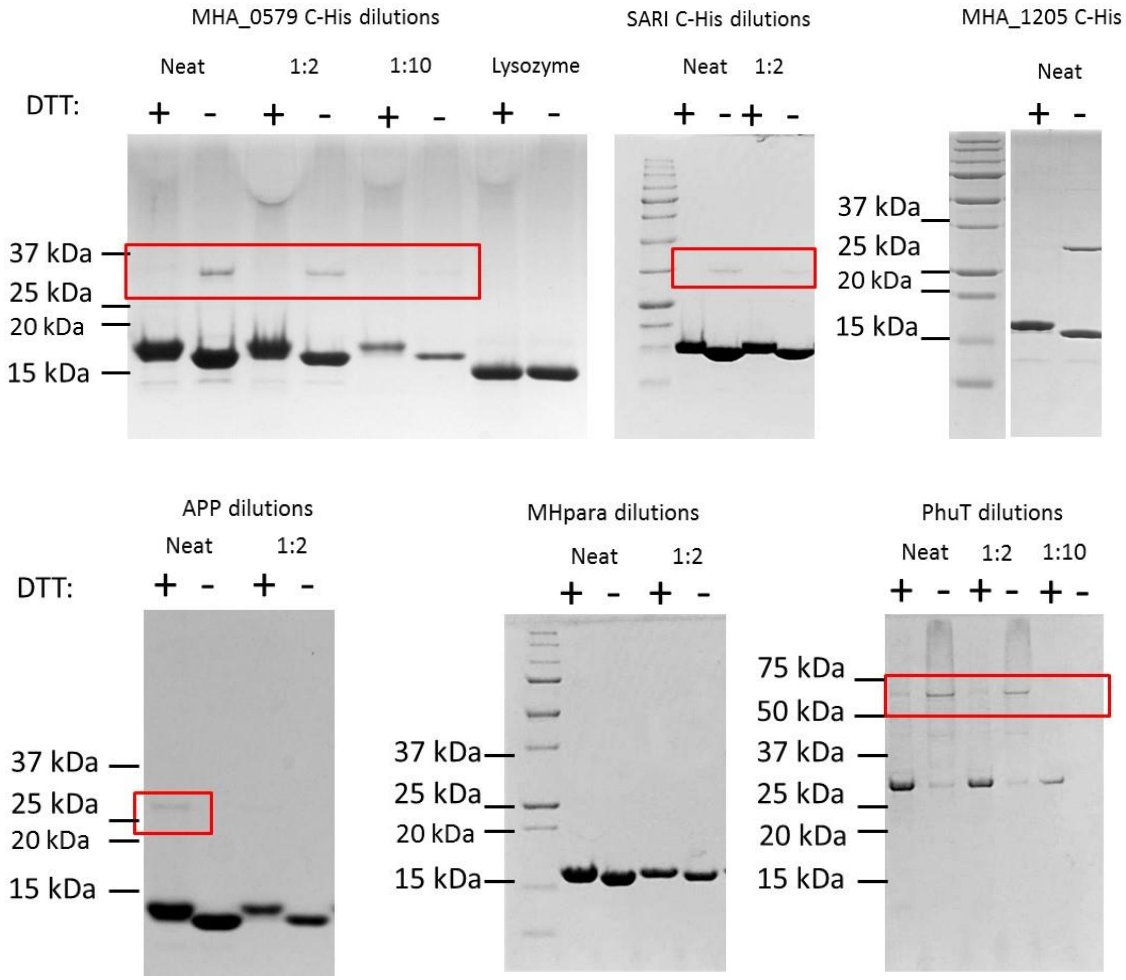


Figure 5.3. Oligomerisation State of Gly1 Homologues Mediated by Disulphide Bonds.

Higher molecular weight bands in samples without DTT (approximately 30-35 kDa) suggested dimer formation of MHA_1205 C-His, MHA_0579 C-His and SARI C-His and PhuT. MHpara and APP Gly1 did not have additional protein bands. Lysozyme was used as a negative control. Proteins boiled for 3 minutes at 100 °C in loading dye with or without DTT prior to loading onto gel then separated by SDS-PAGE. Images are representative of 3 repeats.

To ensure that the higher molecular weight bands seen on the SDS-PAGE gels were in fact the proteins of interest rather than contaminants, SARI C-His and MHA_1205 C-His antisera, specific MHpara and MHA_0579 C-His antibodies and anti-polyhistidine antibodies were used for western blotting (see figure 5.4). APP Gly1 showed a higher molecular weight band in the presence of DTT, which was believed to be a contaminant. The MHA_1205 C-His antisera detected higher molecular weight bands at approximately 30 kDa with both the MHA_1205 C-his specific antisera, suggesting the dimerisation was

real and not contamination of the sample. The MHA_0579 C-His antibody detected the higher MW bands in the sample lanes confirming their identity as MHA_0579 C-His dimers. The band of MHA_0579 C-His monomer with DTT present was at the expected size. In the absence of DTT, the intensity of the monomer band seen on the western blot diminished. This was also observed with APP Gly1, where neither dimers nor monomers were detected at all in the absence of DTT. This could be because of multimer formation at concentrations too low to be detected by the antibody, which decreased the amount of protein in monomer form.

The dimers detected in SARI C-His samples were much more intense than those observed on the SDS-PAGE gel. Again, no dimers were observed in the MHpara western blot in the samples lacking DTT. However, at the highest dilution a higher molecular weight band was observed in the samples with DTT present. This may be a result of the higher concentration of protein used allowing visualisation of these dimers. The higher MW band in the APP Gly1 Coomassie gel was not detected in the western blot, confirming it as a contaminant. The samples lacking DTT ran at a slightly lower molecular weight than those with DTT.

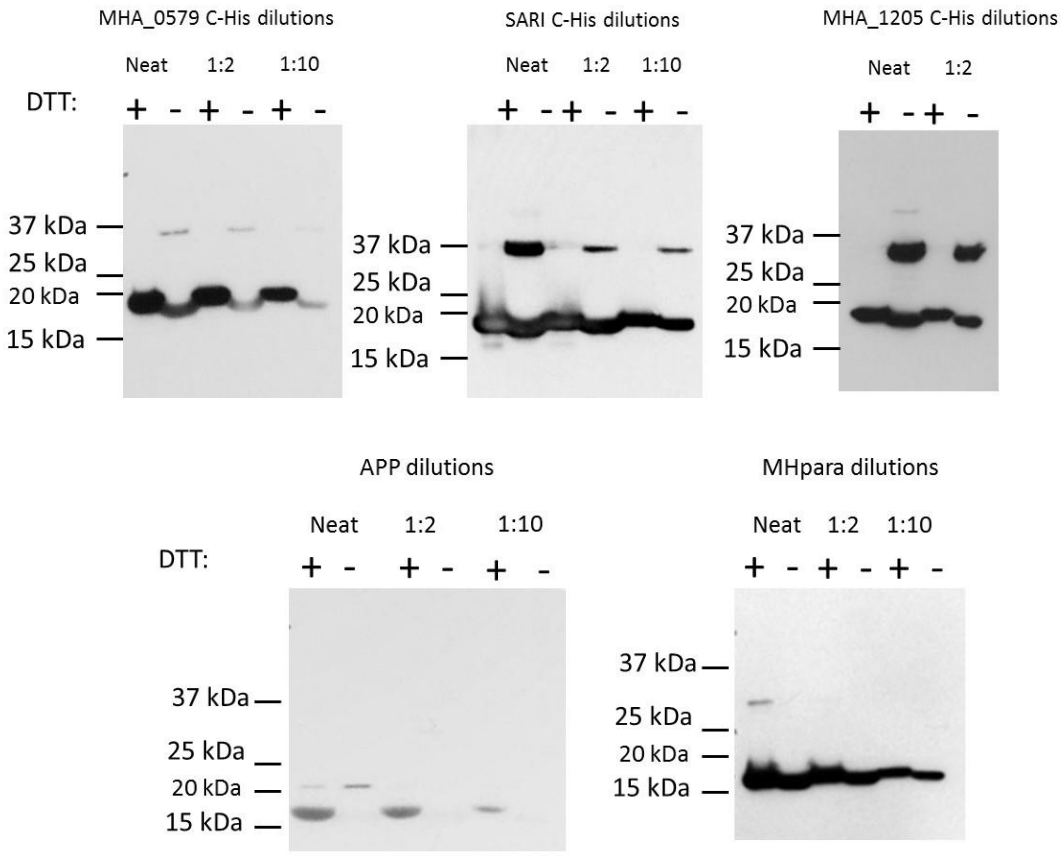


Figure 5.4. Western Blot Confirmation of Protein Dimerisation.

Confirmation of multimer formation using custom made antibodies to each protein. Protein concentration 1 mg/ml in 'Neat' lanes, dilutions made of this. Proteins diluted 1:2 in loading dye and separated by SDS-PAGE. Cognate antibodies or antisera used for each. Proteins boiled for 3 minutes at 100 °C in loading dye with or without DTT prior to loading onto gel. N = 3.

As dimerisation may not necessarily occur through these cysteine disulphide bridges, glutaraldehyde was also used to stabilise any other interactions between these putative homologue monomers by creating cross-links between residues in close proximity. At sufficiently high concentrations, glutaraldehyde can cause artificial binding, so lysozyme was again used as a negative control.

Dimers were observed in the MHA_1205 C-His, MHA_0579 C-His and SARI C-His at a glutaraldehyde concentration lower than that required for lysozyme dimerization, supporting the results observed in the disulphide assay. MHpara and APP Gly1 also showed dimerization at this concentration of glutaraldehyde, suggesting that the dimer interface involves residues other than cysteine (Figure 5.5).

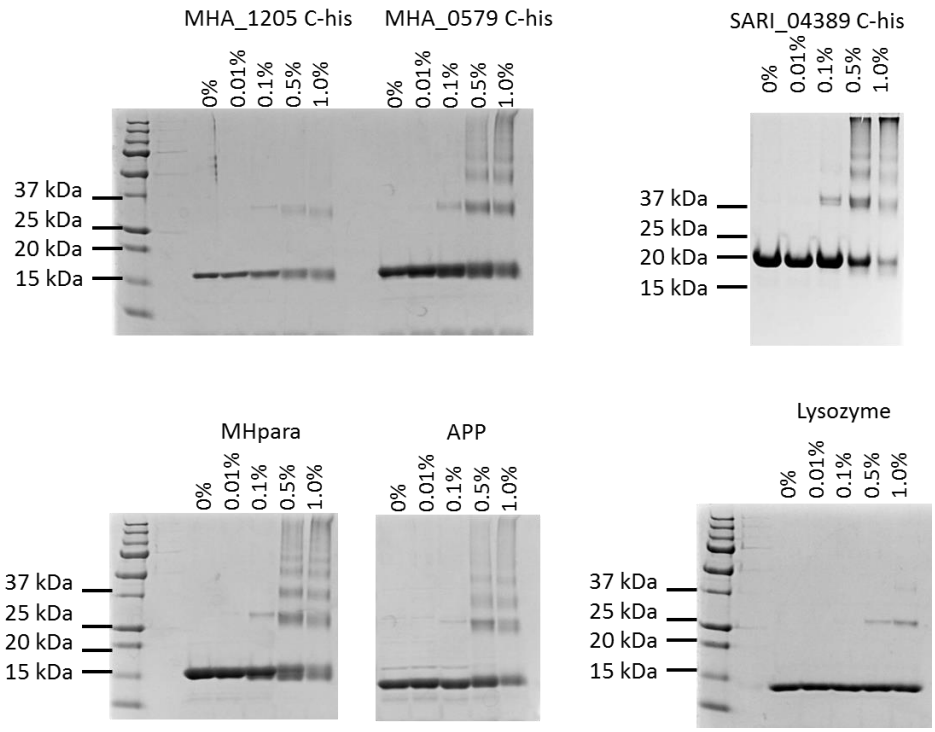


Figure 5.5. Analysis of Multimer Formation of Gly1 Homologues Subjected to Cross Linking.

Multimer formation was analysed using increasing concentrations of glutaraldehyde to stabilise protein interactions. Lysozyme was used as a negative control. Glutaraldehyde percentage stated above each lane. Protein concentration used 1 mg/ml, diluted 1:2 in loading dye containing DTT and separated by SDS-PAGE. Images representative of three repeats.

Confirmation of these higher molecular weight bands as the protein of interest was again performed using protein-specific antibodies (Figure 5.6). DTT was added to the protein samples prior to boiling after glutaraldehyde incubation. The higher percentages of glutaraldehyde (0.5% and 1%) created many species of oligomeric protein throughout the protein solutions analysed, but these were unlikely to be naturally occurring as lysozyme also formed multimers at these concentrations of cross-linker. The Gly1 homologues did, however, form higher oligomers more readily at the high glutaraldehyde concentrations than lysozyme.

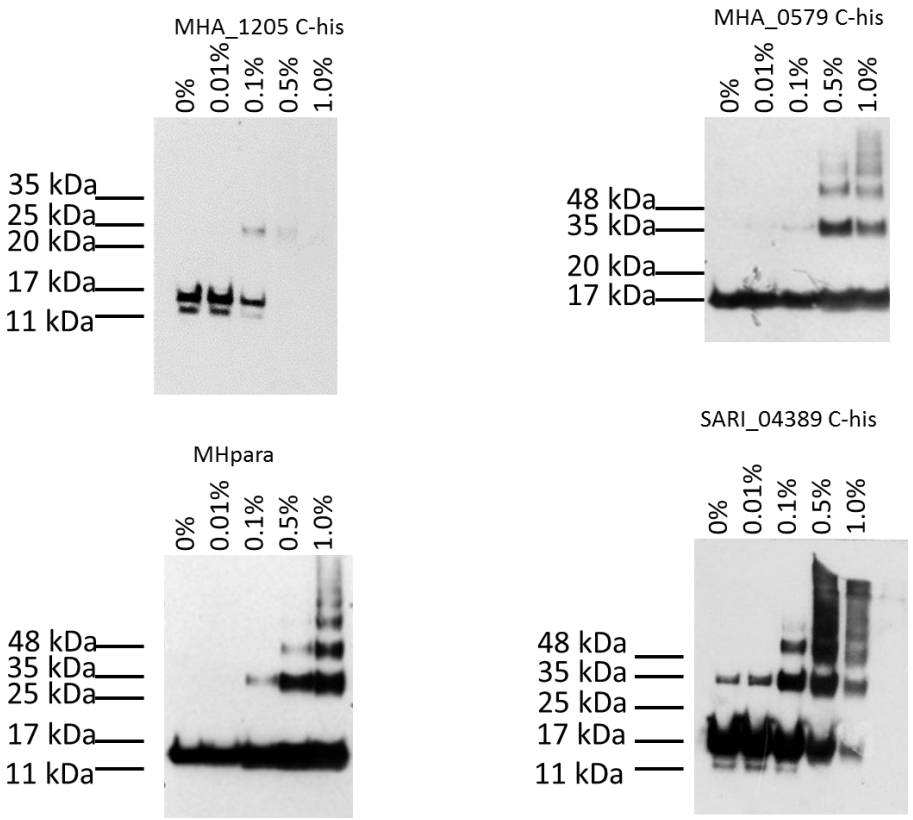


Figure 5.6. Confirmation of Multimer Identity by Western Blot.

Confirmation that the higher MW bands were indeed the protein of interest in multimer form (where antibodies were available). Percentage of glutaraldehyde added shown above each lane. N = 3.

5.2.3 Putative Gly1 Protein Interactions

As discussed in the introduction, bacterial haem uptake systems involve a number of proteins located in different areas of the cell, each with differing functions. To further investigate the function of these putative Gly1 proteins of *M. haemolytica*, the polyhistidine tagged MHA_0579 and MHA_1205 were immobilised to a metal chelate probe on the BLItz (ForteBio). Antibodies were used as positive controls to ensure initial binding or 'loading' of the protein had occurred.

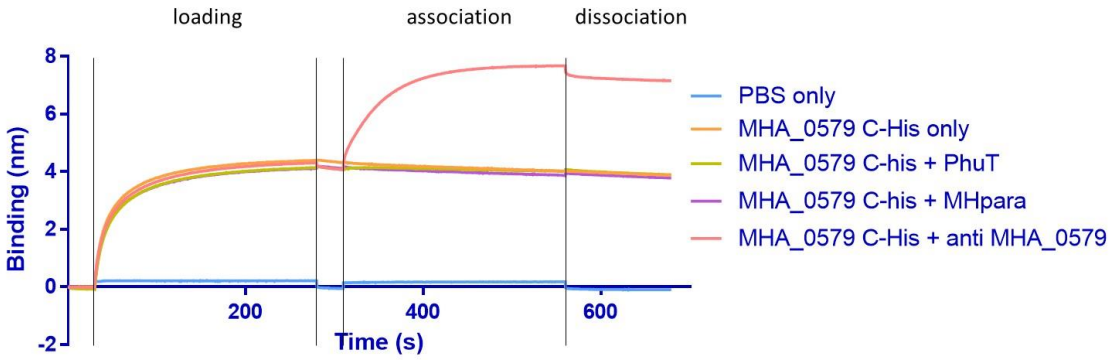


Figure 5.7. MHA_0579 C-His - *M. haemolytica* Protein Interactions.

A Ni-NTA probe was used to immobilise the polyhistidine tagged proteins (loading). Non-tagged proteins were then added to assess interactions (association) and washed (dissociation). No binding was seen between *Mannheimia* proteins. The K_d of anti-MHA_0579 C-his antibodies was estimated to be 4.2 x 10^{-8} M ± 3.13 x 10^{-8} SD, at 0.1mg/ml. PhuT and MHpara alone did not bind to the probe (data not shown). N = 4.

As MHA_0579 was predicted to be an outer membrane protein, it is conceivable that it is involved in the initial cellular binding of haem from exogenous sources such as host hemoproteins or the bacteria's own hemophores for example. To see if MHA_0579 interacted with PhuT or MHpara these proteins were added to the probe after pre-loading with MHA_0579 C-His. No additional binding was witnessed with either of these. The anti-MHA_0579 C-His antibody did display binding to the immobilised protein with a mean K_d of 4.2 x 10^{-8} M ± 3.13 x 10^{-8} SD (Figure 5.7). Quantification of haem binding by this method was not possible at this time.

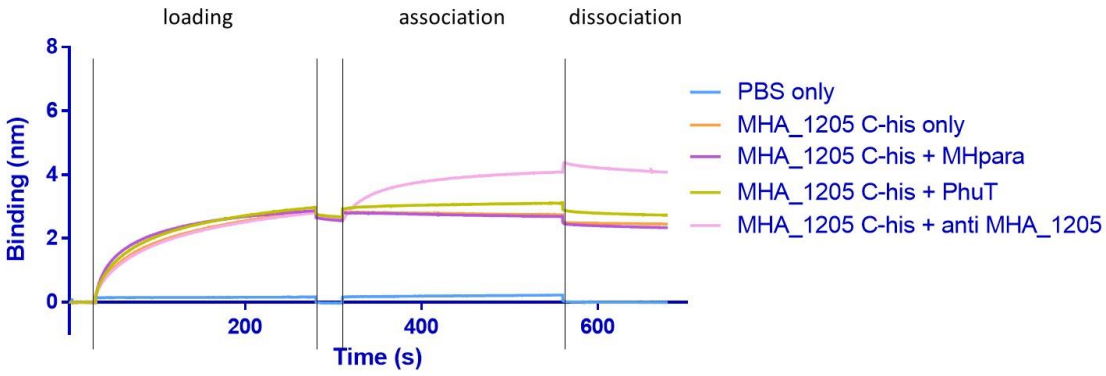


Figure 5.8. MHA_1205 C-His - *M. haemolytica* Protein Interactions.

The K_d of MHA_1205 C-His antisera binding to the K_d was not calculated due to an unknown antibody concentration. PhuT appeared to show a weak interaction with the MHA_1205 C-His giving a mean K_d of 5.6 x 10^{-3} M ± 3.37 x 10^{-3} SD. PhuT and MHpara alone did not bind to the probe (data not shown). N = 3.

MHA_1205 C-His displayed binding to the probe and subsequently to antibodies within the rat antisera. A slight increase in the curve when PhuT was incubated with the pre-loaded probe gave an indication of MHA_1205- PhuT interaction, with an estimated K_d of 5.56 x 10^{-3} M ± 3.37 x 10^{-3} SD on average (Figure 5.8).

5.2.4 MHpara as a lysozyme Inhibitor

The predicted tertiary structure of MHpara was based on that of a lysozyme inhibitor, PliC. These lysozyme inhibitors are an important factor in bypassing the host defence system. To assess whether the MHpara protein was also involved in lysozyme interactions, the two proteins were incubated together in the presence of glutaraldehyde.

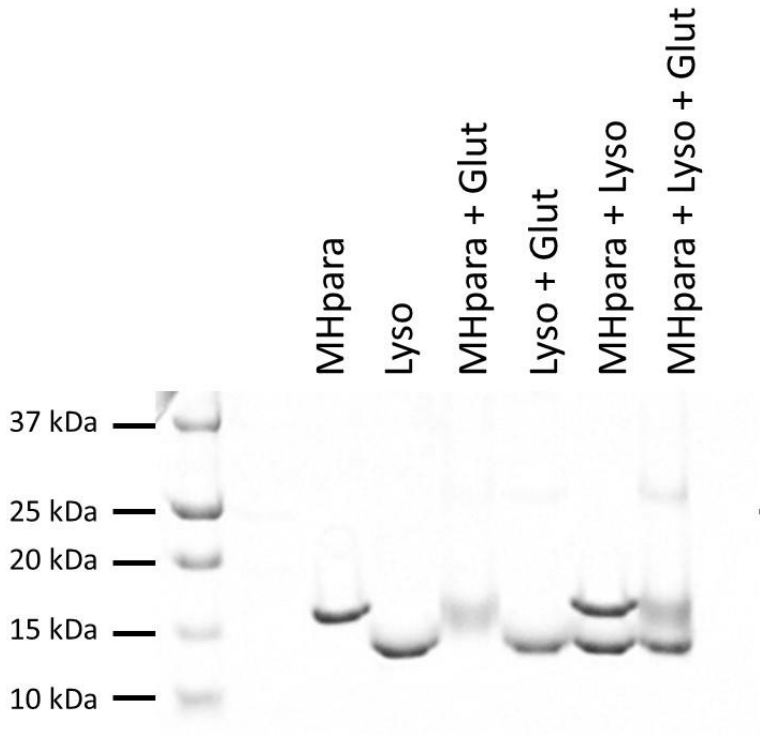


Figure 5.9 Lysozyme Interactions with MHpara

The first two lanes show MHpara and lysozyme alone, followed by the individual proteins incubated with 0.1% glutaraldehyde. The latter two lanes show the proteins incubated together in the absence and presence of glutaraldehyde, respectively. Proteins of 1 mg/ml were used and diluted 1:2 in loading dye and separated by SDS-PAGE. Image representative of three repeats.

The MHpara protein and chicken egg lysozyme both displayed the formation of a small proportion of dimers when incubated with 0.1% glutaraldehyde. The monomers together in the absence of glutaraldehyde did not show any higher molecular weight bands indicative of binding. The higher molecular weight band at approximately 28 kDa was observed when the two proteins were incubated together (final lane, Figure 5.9) to a greater degree than that seen MHpara alone but appears to have migrated through the SDS-PAGE gel at the same speed, suggesting a similar MW.

5.2.5 Interactions of MHpara with Respiratory Tract Tissue

During the course of this project, the annotated name of the MHpara protein on the NCBI database was updated to 'adhesin'. Based on this and the facts that 1) the predicted interactions with haem, hemin or haemoglobin were not seen in any of the

experiments performed, and 2) the lysozyme interaction results did not convincingly suggest binding, the MHpara protein was tested for its ability to interact with respiratory tract tissue of pigs. *M. haemolytica* generally infects cows and sheep, however serotypes A1, A2 and A9 have been isolated from the porcine lung on occasion (Fodor et al., 1999).

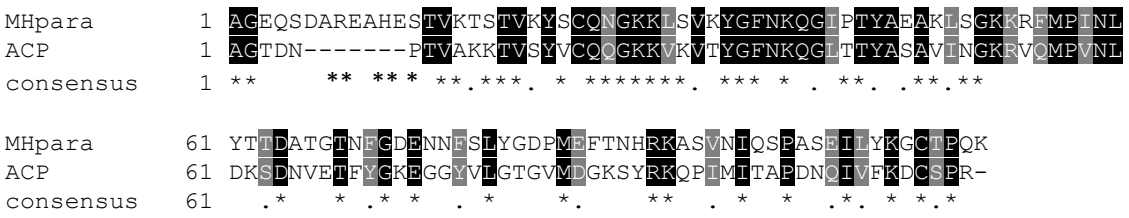


Figure 5.10 Alignment of Mature Protein Sequences of MHpara and ACP *N. meningitidis* Adhesin.

The two amino acid sequences aligned highlighting the identical 42% amino acid identity in the mature protein. Residues shaded black are identical, those shaded grey are similar residues but not identical.

The predicted tertiary structure of ACP was deduced using the Phyre2 server and this predicted structure was then aligned with the solved structure of MHpara (Figure 5.11). The structures had a high degree of similarity displaying the same characteristic antiparallel beta sheets of the Gly1 family. The predicted structure was based on the same protein used for the predicted structure of MHpara, PliC (Chapter 3, Figure 3.16). If these proteins are indeed structurally this similar it is possible that they also share a common function, shown to be adhesion in the case of ACP.

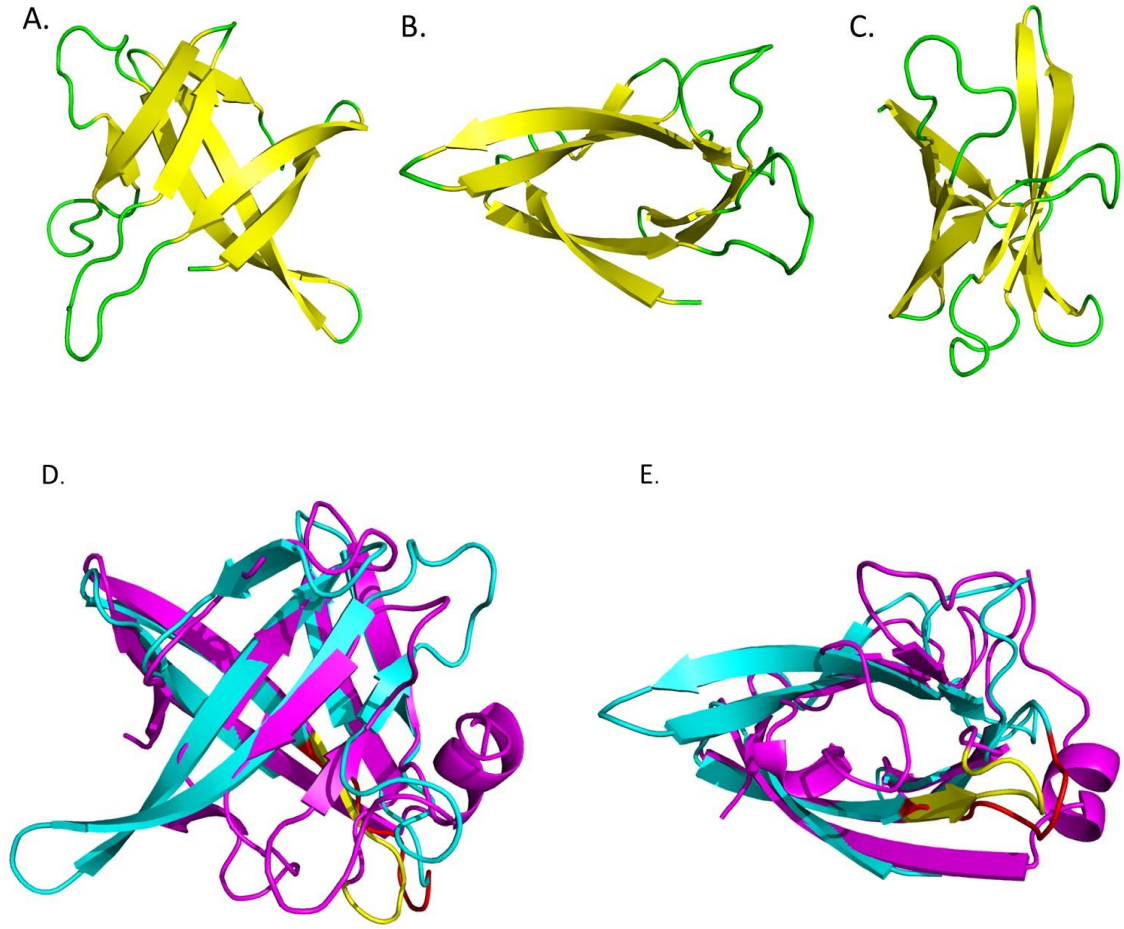


Figure 5.11 Predicted Structure of ACP and Aligned with MHpara.

Panel A, B and C show the predicted structure of ACP generated by the Phyre2 server in three orthogonal views rendered in Pymol. **Panel D** and **E** show ACP (cyan) aligned with MHpara (magenta) from two orthogonal views. The predicted tertiary structure of the *N. meningitidis* adhesin ACP was based on the crystal structure of PliC, periplasmic lysozyme inhibitor of c-type2 lysozyme from *Salmonella typhimurium*. The motif YGFNKQG is highlighted in yellow on MHpara and red on ACP in panels D and E.

M. haemolytica infections are known to take hold in the lower respiratory tract and lung as a result of inhalation of bacteria from the upper respiratory tract following excess growth in an immunocompromised animal. Adhesion is an important virulence factor, and targeting this method of colonisation could potentially prevent infection. Based on this new information, adhesion studies of the MHpara protein were conducted. An initial optimisation assay using a range of dilutions led to the use of MHpara at 20 µg/ml in 1 x PBS for interactions with the porcine respiratory tract transverse sections. The initial concentration of FITC-labelled BSA used in the protocol was in excess at 2 mg/ml. The

labelling of the proteins with FITC was assessed prior to use. Both labelled proteins fluoresced under UV light while the proteins alone did not, showing successful FITC labelling (data not shown). Sections of upper and lower trachea and lung were stained with haematoxylin and eosin (H and E) to highlight the different structures present in the cross-sections (top panel, Figure 5.12). Upper trachea sections were taken from the region inferior to the larynx whereas the lower trachea sections were taken from the area superior to the bronchial bifurcation. This staining clearly shows the epithelial layer on the inside of the trachea section and the cartilage on the outer edge of the upper trachea. The tough texture of the cartilage was problematic during sectioning of the tissue which led to some damage, however, sufficient slides had areas containing an intact epithelium which were focussed on in each of the repeats.

Figure 5.12 shows a section of upper trachea incubated with FITC labelled BSA. DAPI was used to stain the cell nuclei throughout. One small point of green fluorescence was visible in the FITC-BSA control in one repeat, but the trachea was otherwise unstained. When compared to the section incubated with FITC labelled MHpara the difference was very apparent, with widespread staining seen throughout the epithelia and, to a lesser degree, fluorescence in the mucosa. These images are representative of three repeats and support the theory that MHpara is able to adhere to animal epithelial tissue. Unfortunately, no bovine trachea samples were available at the time to allow for inter species comparisons, however, different sections of the porcine respiratory tract were investigated to assess interactions with the different surfaces potentially exposed to the bacteria during infection.

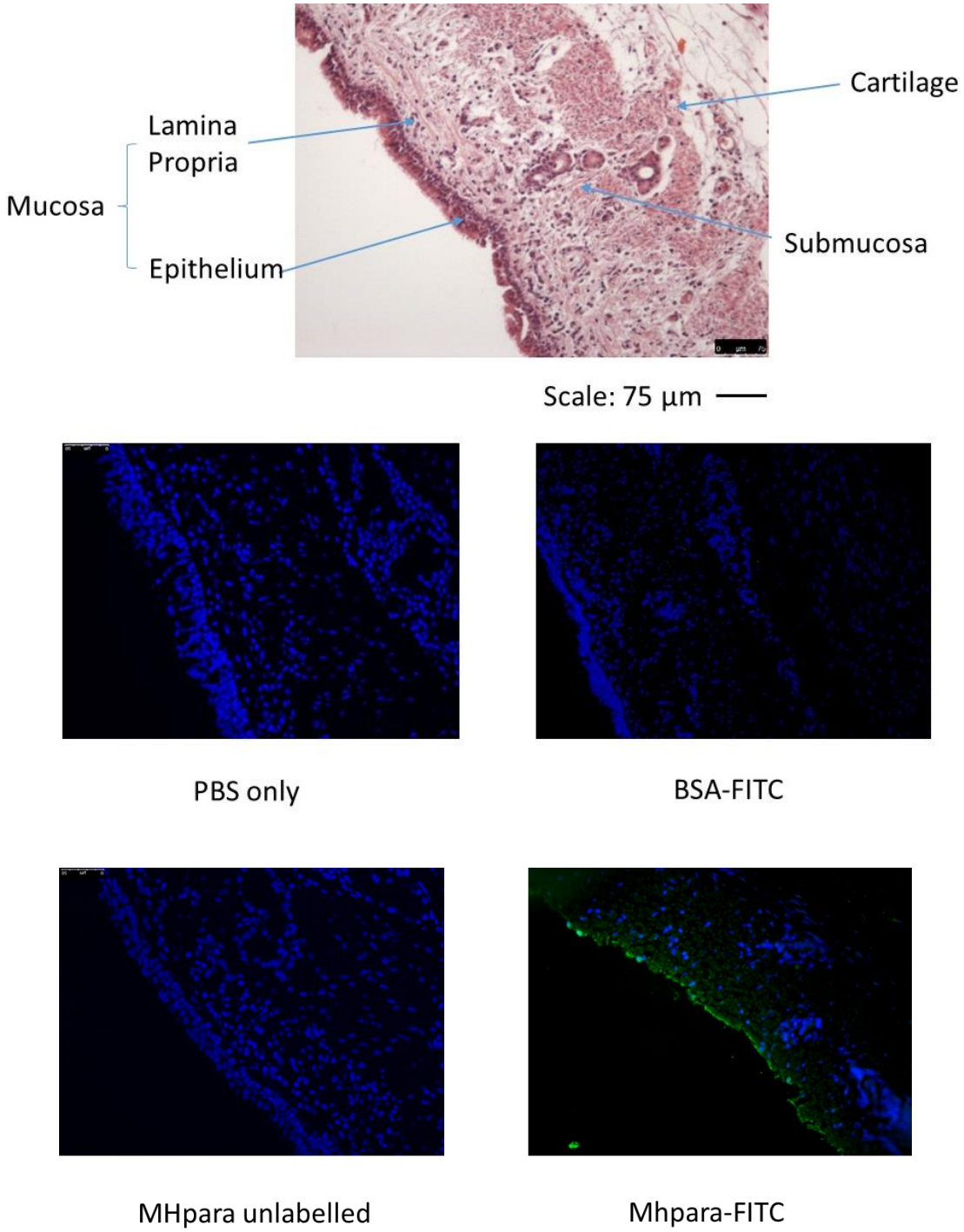
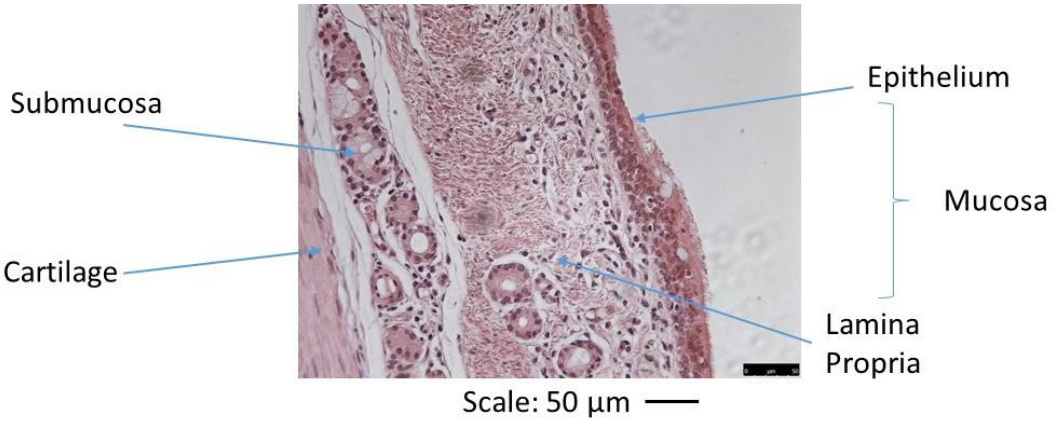
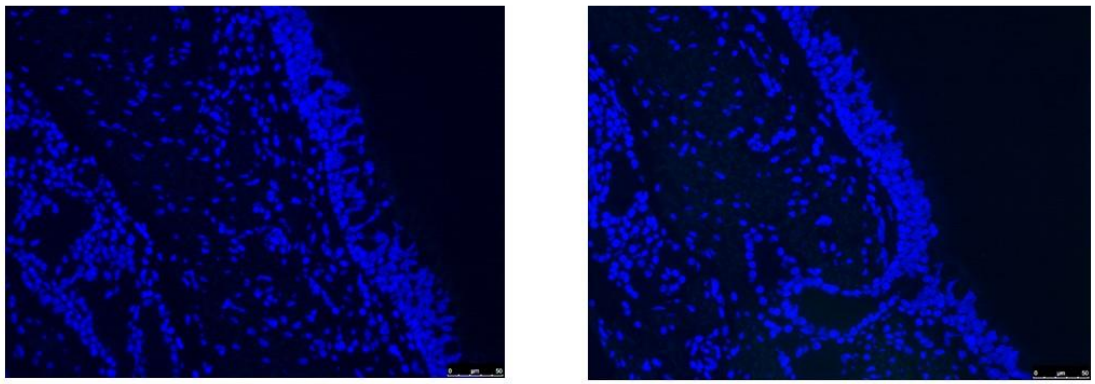


Figure 5.12. MHpara Adhesion to Porcine Upper Trachea Tissue.

The top panel shows an H and E stained section of the upper trachea of the pig, highlighting the epithelial lining. The lower left panel shows the FITC labelled BSA control and the lower right panel, the FITC labelled MHpara lighting up the tissue. The scale bar is 0-75 µm, and the same magnification was used for each section. Repeats were performed on similar sections of tissue and the above images are representative of these. N = 3.

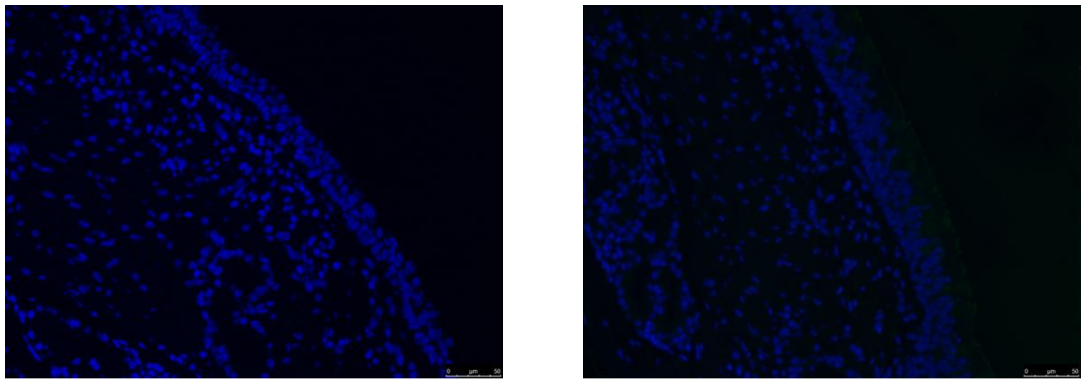
Sections were also taken from the lower respiratory tract of the porcine tissue superior to the bronchus branching. This area of the respiratory system has slightly different morphology to the upper respiratory tract, as the epithelial layer contains fewer goblet cells and the height of the epithelial cells decreases towards the bronchi. The negative controls were the same as for the upper trachea, displaying no detectable green fluorescence at all. The labelled MHpara protein displayed some interactions with this section of the respiratory tract as shown by a low level of green fluorescence. However, this was barely visible on the image, which suggests that binding was not occurring to the same degree (Figure 5.13). No binding was seen between MHpara and the porcine lung tissue (Figure 5.14)





PBS only

BSA-FITC



MHpara unlabelled

MHpara-FITC

Figure 5.13 MHpara Adhesion to Porcine Lower Trachea Tissue.

No MHpara staining was seen in the control slides. Faint green colouring is visible on the outer edge of the epithelial layer of the MHpara-FITC incubated section (to the right hand side of the blue DAPI staining). The scale bars in the lower right corners are 0-50 µM.

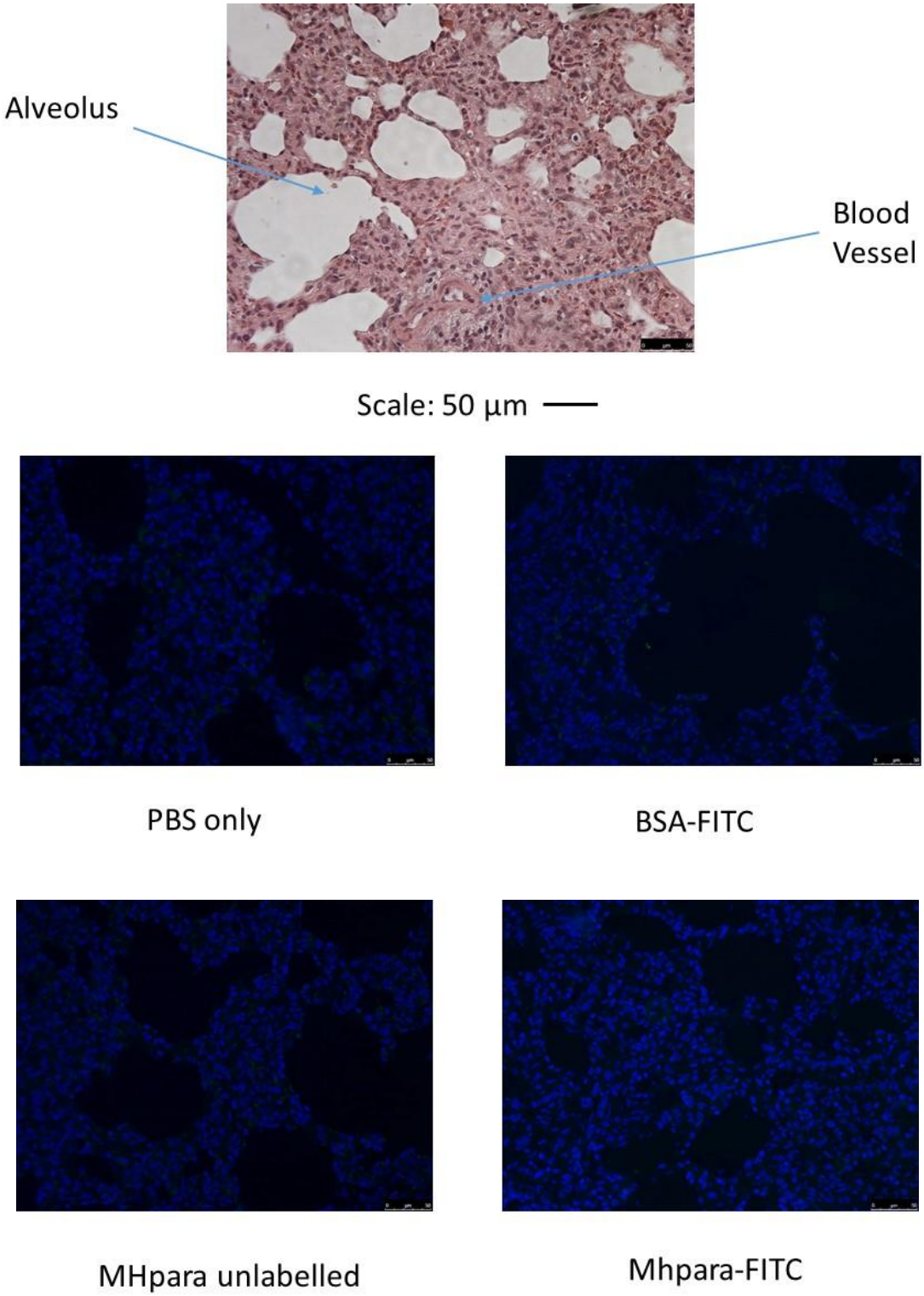


Figure 5.14 MHpara does not Adhere to Porcine Lung Tissue.

No staining was observed after incubation of porcine lung sections with MHpara-FITC protein. N = 3, images are representative of repeats. Scale bars 0- 75 µm.

To analyse the location of fluorescent labelling on the cells, confocal microscopy was used within the departmental core facility. This showed that the MHpara-FITC fluorescence was mainly located on the lumen side of the epithelial layer (Figure 5.15). This supports the theory that MHpara is an adhesin of the upper respiratory tract.

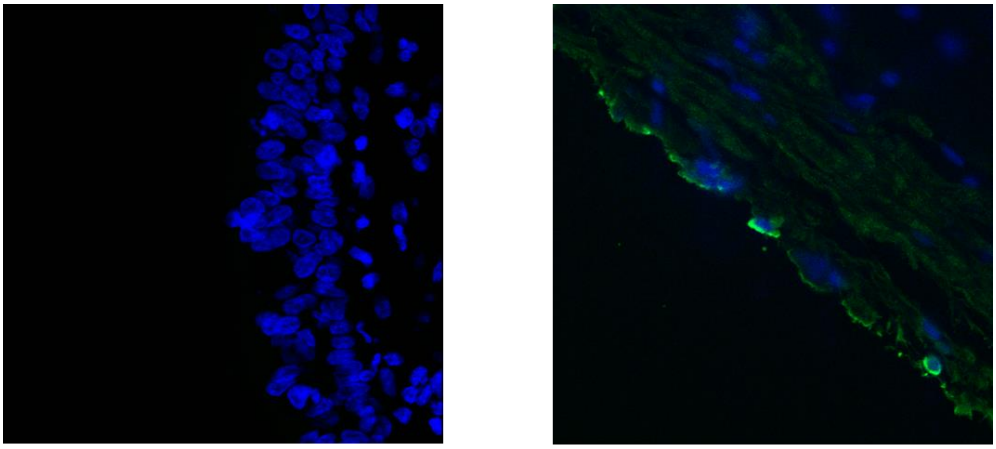


Figure 5.15 Confocal Microscopy of Porcine Respiratory Tract Sections.

Comparison between the PBS negative control and the MHpara-FITC labelled protein. 40x magnification.

5.3 Discussion

This chapter demonstrated that polyclonal Gly1 antisera in general showed no cross-reactivity for the putative Gly1 proteins analysed across species, reducing their potential as broad antibacterial targets. An exception to this was the SARI C-His Gly1 antisera, which demonstrated cross reactivity to MHA_1205, MHA_0579 and to a lesser extent *N. meningitidis* Gly1. Antibodies against the *M. haemolytica* putative Gly1 proteins did not cross react with other *M. haemolytica* Gly1 proteins, possibly due to the limited similarity between the amino acid sequences.

The putative Gly1 homologues displayed different homo-multimer formation to each other, suggesting that they each have different biological functions. Homo-multimers were observed in the absence of DTT with MHA_0579 C-His, MHA_1205 C-His and SARI C-His. Dimers of MHA_1205 C-His were observed at a higher proportion relative to those of MHA_0579 C-His when compared to the intensity of the monomer band, which could suggest that it forms these dimers more readily. SARI C-His also formed dimers although these were more evident in the western blot than the SDS-PAGE gel, which is possibly due to the high sensitivity of the antibodies to the protein present compared to the Coomassie stain. Disulphide bonds between cysteine residues can be intermolecular or intramolecular, and breaking these intramolecular bonds possibly led to the protein opening out and passing through the gel more slowly than the relatively compact bonded protein.

The differences in dimerisation seen between the proteins could be a result of the location of the two cysteines in the processed protein tertiary structure, or simply the lower percentage identity to each other leading to differing interactions, suggesting again that these proteins behave in a different manner. The high proportion of dimers formed suggests that dimerisation may be important for the function of MHA_1205. Due to the lack of structural information about these proteins at the time of writing we cannot judge how these disulphide bridges may form in terms of amino acid residue proximity, and therefore how the monomers may interact with each other. This interaction could provide clues as to the mechanism of haem binding of these proteins, such as creation of binding pockets, and also further information of their role within this

nutrient uptake system. Czjzek *et al* discovered a dimeric form of the HasA hemophore in *S. marcescens* where each monomer bound haem in a 1:1 ratio (Czjzek et al., 2007). The dimer was a proposed haem reservoir as it had to revert to the monomer form to interact with the receptor, HasR, to deposit the haem. This dimerisation was found to occur by domain swapping, which cannot be ruled out as the dimerisation mechanism between the putative Gly1 homologues as the disulphide bridges could form as a result of this movement.

Neither MHpara nor APP proteins formed dimers in the absence of DTT, and as neither of these proteins displayed any interactions with haem this provides more evidence to suggest that they are not closely related to the other putative Gly1 homologues in terms of behaviour. Transient interactions are possible, but the glutaraldehyde cross linking should have accounted for this and again this showed the protein was negative for multimers at the optimal glutaraldehyde concentration. PhuT on the other hand appeared to form dimers very readily, which has been observed with other Gram-negative bacterial haem PBPs even when interacting with haem in a 1:1 ratio, providing further support for PhuT as a haem PBP (Mattle et al., 2010, Ho et al., 2007). With a high initial protein concentration, dimers of MHpara were observed as seen with 0.1% glutaraldehyde, but this was a low proportion compared to the monomers present.

A limited number of solved crystal structures of haemoglobin and hemin outer membrane receptors have been published to date, and all of these are in monomer form. It is interesting therefore that the predicted OM protein MHA_0579 C-His, which has been shown to interact with hemin, was also seen to form dimers. The dimers were only visible on SDS-PAGE gels when loading higher concentrations of protein. It is possible that the lower proportion of dimers seen were an artefact of the experimental conditions and the MHA_0579 protein does not form dimers to a physiologically relevant extent *in vivo*. It would be interesting to compare the production of this protein under different growth conditions as a large up-regulation of transcription could provide clues as to the likelihood of dimerisation to a meaningful extent *in vivo*.

As shown by the bipartite OM receptor in the lactoferrin uptake system, both proteins can bind iron to enable uptake into the cell but have different roles in the uptake process. The predicted cellular locations of MHA_1205 and MHA_0579 possibly suggest

different haem uptake roles. It is possible that the identified potential Gly1 homologues and paralogues work within the haem uptake system but have different roles. Sprencel *et al* identified interactions between the siderophore, OM receptor and PBP of the Fep iron uptake system of *E. coli* (Sprencel et al., 2000). The predicted cellular location of the proteins was the outer membrane for MHA_0579, and the periplasm for MHA_1205 and PhuT. The cellular location of MHpara was not conclusively predicted (Chapter 3, Table 3.4). This again points to different roles within the cell. It is unknown whether PhuT is the PBP used by the system involving any of the *M. haemolytica* Gly1's; one of these Gly1 homologues could be a haem PBP itself. The weak interaction between PhuT and MHA_1205 C-His provides some evidence to further support the theory that MHA_1205 is involved in the haem uptake pathway as a whole, although it does not provide any indication on the particular role within the system. In the presence of an interaction partner such as hemin this binding affinity may change.

Outer membrane proteins tend to be larger than 20 kDa in size, however, the MHA_0579 protein was predicted to be around 13.7 kDa in its processed form. We know that transferrin and lactoferrin outer membrane receptors are bipartite, formed of two distinct monomers, each of around 75- 100 kDa. It has also been demonstrated that the hemophores, e.g. HasA, interact with an outer membrane receptor such as HasR. The results of the interaction studies performed in this project suggested that the Gly1 proteins of interest did not interact with each other. This does not rule out the possibility that they function as part of different haem uptake systems, as some bacteria are known to have multiple acquisition methods as seen with the multiple haemoglobin receptors in *N. meningitidis* (Tauseef et al., 2011). It is also possible that the proteins require haem to interact, as this can induce conformational changes. Interestingly, PhuT was observed to interact with MHA_1205 C-His, which may indicate MHA_1205 as having a chaperone type role in the periplasm.

In addition to analysing individual interactions with known proteins using the BLItz technology, use of a bacterial two-hybrid system to screen a greater number of potential binding partners for these proteins would provide further insight. Again, confocal microscopy using specific antibodies would provide important information to clarify the

cellular location of the individual proteins, and from this a greater idea of the function of said proteins.

The comparison of the solved and predicted tertiary structures of MHpara highlighted the structure as similar to that of PliC from *S. typhimurium*, a lysozyme inhibitor. This protein shares a high degree of similarity in terms of the overall structure with MliC of *P. aeruginosa* despite only sharing 27% identity in the aligned regions (Leysen et al., 2011). Based on this, glutaraldehyde was used to analyse any interactions between MHpara and chicken egg lysozyme. The results of this did not show binding but were not fully conclusive, with a higher MW band detected but no way of confirming the presence of lysozyme in the dimer at this point. If this higher molecular weight protein band was the product of lysozyme and MHpara, it would potentially be a lower MW than that seen faintly in the MHpara alone lane than a MHpara dimer depending upon the manner of interaction. This did not appear to be the case, suggesting that it was not the product of MHpara and lysozyme interacting with each other. Chicken egg lysozyme and bovine lysozyme are the same class (C class) but only share around 20% sequence identity, which could present a reason for the lack of perceived interaction. A lysozyme antibody could also be used to check for protein co-localisation.

Further analysis revealed that the MHpara protein had an entirely different function. In *N. meningitidis*, an adhesin of similar size to the MHpara protein was described by Hung *et al* (Hung et al., 2013). This highly conserved 124 a.a., 13.3 kDa Adhesin Complex Protein (ACP) is the product of gene NMB20925 in *Neisseria meningitidis* strain MC58 - the same strain used in the Sayers lab for Gly1 investigations. Meningococcal cells expressing ACP were shown to adhere *in vitro* to Chang, Hep2, HUVEC and meningioma cells significantly more than ACP knock out strains, with the highest level of adhesion observed to epithelial cell lines. This protein has a predicted 22 a.a N-terminal signal peptide and shares approximately 25.6% identity with the unprocessed MHpara protein, and 42% when comparing the mature proteins (see Figure 5.10). The ACP protein also demonstrated significant complement-mediated bactericidal activity in the same study.

This protein was found in the OM and OM vesicles, and was upregulated in conditions of iron paucity (Grifantini et al., 2003). Antibodies to ACP promoted complement mediated killing and one of three types of ACP identified was found in all investigated

strains (Hung et al., 2013). Adhesion studies were therefore conducted with the MHpara protein.

Porcine respiratory tract was used, which bears similarity to bovine and ovine respiratory tract in terms of epithelial cell types. It has been shown that *M. haemolytica* displays species specificity in terms of colonisation and pathogenicity. However, the aims of this study were to assess protein- epithelial cell binding. The host species specificity is thought to be due to differences in bacterial OM proteins and recognition by alveolar macrophages in terms of bovine and ovine infection (Hounsome et al., 2011). Adhesion of the MHpara protein was observed in the trachea but not in the lung. The main differences between these sections of the respiratory tract are the epithelial lining and the presence of either mucous or pulmonary surfactant in the trachea and lung respectively. The binding seen in vitro of MHpara to the epithelial cells of the trachea occurred on formaldehyde preserved tissue. It is likely that the mucous layer normally present on airway epithelial cells was lost during the tissue fixation process which may have influenced the results, therefore it would be beneficial to do further work using air- liquid interface (ALI) techniques to recreate an environment closer to that occurring *in vivo* to deduce whether or not this binding still occurs, or in an appropriate animal model.

The most reliable method of assessing *M. haemolytica* infection *in vivo* is using bovine models of disease, however, use of this method is limited by cost. Mouse models of disease have had variable success (Gatto et al., 2006, Jian et al., 1991). A 'bovinised' mouse model of *M. haemolytica* infection has recently been described (Stellari et al., 2014) although this concentrates on bovine IL-8 expression within the mouse. Issues with the murine model included rapid clearance of bacteria, and different types of lung lesions to those observed in larger mammals. An animal model of infection would be useful to observe the differences between infection with wild type and MHpara null mutant strains.

While small animal disease models are available for *A. pleuropneumoniae* infection (Bhatia et al., 1991), *Katribe et al* published a study showing that *S. enterica arizonae* strains IIIa and IIIb were unable to infect the murine host, possibly due to the preference

for cold blooded animals, and therefore were not a viable infection model to study knock out strains of these organisms in the future (Katribe et al., 2009, Chiang et al., 2009).

In conclusion, MHpara appears to be a bacterial adhesin based on the respiratory tract interaction study. The other *M. haemolytica* putative Gly1 proteins do not interact with each other. They were, however, shown to form dimers by disulphide bridges where the MHpara and APP proteins did not, and could comprise different parts of a haem uptake system in *M. haemolytica, A. pleuropneumoniae* and *S. enterica arizonae*. While MHA_1205 interacted with PhuT, the significance of this needs further investigation. The function of APP Gly1 remains elusive.

6. Screening BRD Bacterial Isolates for Gly1 homologues

6.1 Introduction

The genome sequences of *M. haemolytica* available on the NCBI database are of known invasive isolates- A1, A2 and A6. This is a limited selection considering that 12 strains have been identified to date after a number of reclassification events within the genus as a whole. These three strains are, however, the most clinically significant as they are the most common serotypes isolated from diseased animals, with A1 making up the majority of disease cases. The NCBI database also describes an A1/A6 cross-reactive isolate; a proposed novel strain of *M. haemolytica* taken from the spleen of a pneumonic white tail deer (Lawrence et al., 2014).

The prevalence of different *M. haemolytica* isolates has been studied in countries such as Hungary, Denmark, USA, and Canada. Of these, the serotypes of strains found in cattle, sheep and goats have been deduced where possible, although some remain untypable. Similar strains are predominant in different countries, for example in both Northern Ireland and Hungary strains A1 and A2 were most commonly isolated, comprising approximately 60% and 20% of isolates, respectively (Fodor et al., 1999, Ball et al., 1993). This is a trend repeated in many regional studies. Pasteurellosis in sheep is most commonly caused by serotype A2, but the ovine host is also colonised by a much wider range of *M. haemolytica* serotypes than cattle or goats and A6 and A9 have also been found to contribute to the ovine bacterial fauna in high percentages (Odendaal and Henton, 1995). There has been an increase in the number of the 'less common' serotypes being isolated globally, although it is unclear if this is as a result of a wider research interest in BRD or to an actual increase in the prevalence of these strains. The serotypes differ mainly in the makeup of their capsule, outer membrane proteins and LPS types (McCluskey et al., 1994).

Sequence analysis for this project showed that one or more of the potential Gly1 homologues of *M. haemolytica* was found in 16 out of the 20 available genomes on the NCBI database. The paralogue was identified in all but one; the presumed *M. haemolytica* strain isolated from deer that displays a clear variation from others characterised to date. The only strain of *A. pleuropneumoniae* that had a potential Gly1

was 6 FEMØ out of the 20 strains on the database. In *Salmonella*, of all the hundreds of genome sequences currently available, only one serotype had a possible candidate as a Gly1 homologue- *S. enterica arizonae (IIIa).*

For Gly1 proteins to be a plausible target for a vaccine they would have to be expressed in a high percentage of the known pathogenic serotypes and their various isolates to allow for Gly1 specific antibodies to direct killing of the cells in question. The aims of this chapter are to assess the prevalence of the Gly1 genes as potential vaccine targets by screening a selection of *Mannheimia* field isolates. Therefore the aim of the work described in concentrated on *M. haemolytica*, aiming to investigate the presence of Gly1 proteins in a variety of isolates from England. Isolates were obtained from the RVC and AHVLA, and their identity investigated by culture techniques and 16S gene analysis. Primers were designed to detect Gly1-like genes within these isolates, and any sequences obtained were analysed.

6.2 Results

6.2.1 Isolate Identification

Confirmation of the identification of donated isolates as *M. haemolytica* was carried out prior to screening. As the samples were mainly bovine isolates obtained from pneumonic animals, there was a risk of contamination with other *Pasteurellaceae* species such as others implicated in BRD complex. Original identification methods of the individual strains (where labelled) were not known due to the age of the samples received from the Royal Veterinary College (RVC) and uncertainty over the methods used on the samples from the Animal Health and Veterinary Laboratories Agency (AHVLA). Initial growth on horse blood agar plates was assessed alongside growth on MacConkey agar plates. *M. haemolytica* colonies appear small, circular and grey on blood agar plates, and as the name suggests display some weak beta-haemolytica activity caused by leukotoxin, which is especially prominent on sheep blood agar. When streaked on MacConkey agar plates *M. haemolytica* either does not grow or forms very small circular pink colonies after 48 hours. This is due to the ability to ferment lactose, albeit weak. Lactose fermenting bacteria internalise the neutral red pH sensitive dye found in the MacConkey agar and as a result turn pink.

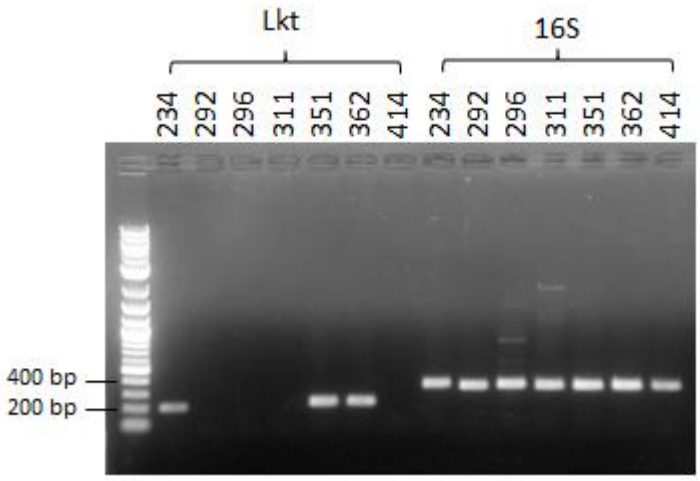


Figure 6.1. Confirmation of Identity of Isolates by 16S RNA PCR Screening.

16S primers (see Primer Table 2.4) were used to show that sufficient genomic DNA was present in the colony boilate, and confirm PCR success. PCR with Lkt primers showed that isolates 234, 351 and 362 were positive for *M. haemolytica*, while 292, 296, 311 and 414 were not. Samples separated by agarose gel electrophoresis. This was repeated with all RVC isolates. Negative controls lacking either the genomic DNA or primers were also used throughout (data not shown).

Based on these results, the AHVLA isolates 234, 235, 351, 362 and 428 were positively identified as *Mannheimia* species. Strains 296 and 311 grew as large, yellow, *E. coli*- like colonies, suggesting either contamination of the samples received or incorrect samples. Isolate 411 did not grow at all and 414 did not grow in liquid culture preventing the creation of freezer stocks for further growth-based analysis. This initial screening was repeated for the RVC isolates, of which all but A7, A10 and MH9 A2 grew on MacConkey agar appearing as small pink colonies.

Further identification procedures were based on the paper by (Alexander et al., 2008). Universal 16S primers were used to confirm PCR success, and *M. haemolytica* Lkt-specific primers were used to identify the *M. haemolytica* positive samples (Figure 6.1). PCR was performed on the extracted isolate DNA and products were purified and sent for sequencing allowing for sequence comparison. The 16S sequencing for the strain named A10 did not align with the *M. haemolytica* 16S sequences in the database. The bacterial isolate was named A10 when it was isolated and characterised in 1975, prior to a number of reclassification events. The strain was subsequently called T10 when the

Pasteurella species was split into A and T biovars, followed by further reclassification as *B. trehalosi*. The isolate named A10 was confirmed by a BLASTn search of the amplified gene sequence to be this species, which showed sequence identity with a 303 bp fragment of *B. trehalosi* 16S. Using this method, isolate p28 was identified as *M. varigena*. Isolate 414 16S sequencing shared 100% identity with *E. coli*, a result that was confirmed by the 23S sequencing. The 23S primers were also used as a further aid to identification of the strains, and all 23S sequences supported the 16S results. A summary of the isolate identification results is displayed in Table 6.1.

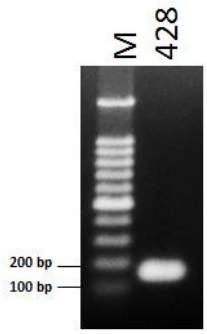
	16S	Lkt	Lkt2	23S	Growth on MacConkey	*M. haemolytica*?
A10	✓*	-	-	✓	N	*B. trehalosi*
A12	✓	✓	-	✓	Y*	Y
P28	✓*	-	-	✓	Y*	*M. varigena*
A1	✓	✓	-	✓	Y*	Y
A7	✓	✓	-	✓	N	Y
A9	✓	✓	-	✓	Y*	Y
CO2392	✓	✓	-	✓	Y*	Y
CO0600	✓	✓	-	✓	Y*	Y
CO4178	✓	✓	-	✓	Y*	Y
CO3983	✓	✓	-	✓	Y*	Y
CO3916	✓	✓	-	✓	Y*	Y
Gilly	✓	✓	-	✓	Y*	Y
CO0353	✓	✓	-	✓	Y*	Y
MH4 A8	✓	✓	-	✓	Y*	Y
MH8 A5	✓	✓	-	✓	Y*	Y
MH9 A2	✓	✓	-	✓	N	Y
A6	✓	✓	-	✓	Y*	Y
CO2008	✓	✓	-	✓	Y*	Y
CO1250	✓	✓	-	✓	Y*	Y
234	✓	✓	-	✓	Y*	Y
235	✓	✓	-	✓	Y	Y
292	✓*	-	-	✓	Y	*Acinetobacter*
296	✓*	-	-	✓	Y	mix
311	✓*	-	-	✓	Y	*E. coli*
351	✓	✓	-	✓	N	Y
362	✓	✓	-	✓	N	Y
411	N/A	N/A	N/A	N/A	N/A	?
414	✓*	-	-	✓	N	*E. coli*
428	✓	✓	✓	✓	N	*M. glucosida*

Table 6.1 Isolate Identification Summary.

A summary of the genes present in each isolate according to PCR and growth on MacConkey agar plates. The 16S* refers to genes that were amplified but do not correspond to *Mannheimia* 16S genes in the NCBI database. If small pink colonies grew on MacConkey agar, Y* is denoted. N/A here indicates that the 411 strain did not grow at all, even after repeated attempts with fresh stocks. 296 was a contaminated stock of *M. haemolytica.*

Sequencing of the gene fragments amplified by the 16S, Lkt, Lkt2 and 23S primers assisted with further identification of the isolates. The primer set designed to amplify the Lkt2 fragment of the leukotoxin gene in *Mannheimia* species was used to identify

M. glucosida. Isolates 234, 235, 351 and 362 of the AHVLA set were determined to be *M. haemolytica* based on BLASTn searches of the sequenced PCR products, as were all the RVC isolates that were positive for Lkt. Isolate 428 from the AHVLA was deemed to be *M. glucosida* as 16S sequencing reported matches to both *M. haemolytica* and *M. glucosida*, Lkt sequencing suggested *Mannheimia* species but the presence of the Lkt2 gene fragment confirmed *M. glucosida* was present within this sample (Figure 6.2).




```
Lkt2_428R     1 --------------------------CGTAGTCGATCTTCTTCCGATGA-TCGTTAAATTCGGTTTTAGATA
LktABCDop     1 ATTGCTCCTCAATTAAGTGTTTAATACGTAGTCGATCTTCTTCCGATGAATCGTTAAATTCCGTTTTAGATA
Lkt2_428F     1 ATTGCTCCTCAATTAAGTGTTTAATACGTAGTCGATCTTCTTCCGATGAATCGTTAAATTCGGTTTTAGATA
consensus     1 ........................**********************.***********.**********

Lkt2_428R    73 AATCAATCACCGGCAATGACTCGTTTTCAATGGCAGTCAGTAATGTTTGGTAGCGATAGTT
LktABCDop    73 AATCAATCACCGGCAATGACTCGTTTTCAATGGCAGTCAGTAATGTTTGGTAGCGATAGTT
Lkt2_428F    73 AATCAATCACCGGCAATGACTCGTTT-CAATGGCAGTCAGT--------------------
consensus    73 *************************.***
```

Figure 6.2. Mannheimia Isolate Identification.

The Lkt2 primers amplified a fragment of approximately 160 bp in isolate 428. Sequencing results used in a BLASTn search matched the *Mannheimia glucosida* strain PH290 lktCABD operon, aligned with the 428 Lkt2 sequences above (accession number AF314517).

6.2.2 PCR Screening for Gly1 Homologues

Once isolates had been identified as *Mannheimia* strains, PCR sequencing was performed with sets of gene-specific primers designed using information from the NCBI database genomes (Chapter 2, Table 2.4). This method was chosen as initially only the *Mannheimia* isolates donated by the AHVLA were available for screening. The RVC isolates were then also screened using this method with the same primer sets for continuity. A summary of the screening results can be seen in Table 6.2.

	A (MHA_0579)	**D**	**H** (MHpara)	**J** (COI_1584)	**L**	**M**	**PhuT**	**HmbR**
A12	✓	COI_1584	✓	?	✓	MHA_1205	✓	✓
A1	✓	✓	✓	?	✓	COI_1584	✓	✓
A7	-	-	✓	-	-	-	✓	✓
A9	✓	-	✓	?	✓	COI_1584	✓	✓
CO2392	-	-	*	-	-	-	-	✓
CO0600	-	✓	✓ I to F	?	✓	MHA_1205	✓	✓
CO4178	✓	✓	✓	?	MHA_1205	MHA_1205	✓	✓
CO3983	✓	✓	-	?	MHA_1205	MHA_1205	✓	✓
CO3916	-	-	✓	-	-	-	✓	✓
Gilly	✓	✓	✓	?	MHA_1205	MHA_1205	✓	✓
CO0353	✓	✓	-	?	MHA_1205	MHA_1205	✓	✓
MH4 A8	*	COI_1584	✓	?	✓	COI_1584	✓	✓
MH8 A5	✓	-	✓	?	✓	COI_1584	✓	✓
MH9 A2	-	-	✓	-	-	-	✓	✓
A6	✓	✓	-	?	MHA_1205	MHA_1205	✓	✓
CO2008	✓	✓	✓	?	MHA_1205	MHA_1205	-	✓
CO1250	✓	✓	✓	?	MHA_1205	MHA_1205	✓	✓
234	-	COI_1584	✓	*	✓	* COI_1584	✓	✓
235	✓	-	-	-	✓	-	✓	✓
296	-	COI_1584	-	-	-	-	-	-
351	-	COI_1584	✓	✓	✓	COI_1584	✓	✓
362	-	-	*	-	MHA_1205	-	✓	✓
428	-	-	-	-	MHA_1205	-	-	-

Table 6.2 Summary of Positive PCR Screening for *gly1* Genes.

Results of screening donated bacterial isolates for genes of interest. Full sequencing of the HmbR PCR products was not performed. Partial sequencing confirmed the gene identity. Ticks indicate the presence of the expected gene with each particular primer set, alternative genes found with each primer set are stated in the corresponding box, and those with slight differences are indicated with an asterisk. D, L and M sets designed to detect MHA_1205, COI_1584 and L280. Sets A-M were designed to detect homologues and the paralogue.

Isolate 428 was identified as *Mannheimia glucosida* rather than *M. haemolytica* which could explain the lack of positive results for *gly1* genes such as MHpara and HmbR. *M. glucosida* is found primarily in sheep and along with *M. haemolytica* is a cause of ovine mastitis.

No whole *M. glucosida* genomes have been uploaded to the NCBI database at present (July 2015), and all *M. glucosida* genes that have been uploaded have been the target of a specific study e.g. that of the transferrin binding receptor (Lee and Davies, 2011). This study also investigated the process of horizontal gene transfer between *M. haemolytica*, *M. glucosida* and *P. trehalosi*, although the transferrin binding protein '*tbpA*' gene is believed to be conserved due the evolution of the species from a common ancestor potentially supporting the theory that putative *gly1* genes are present in these *pasteurellaceae* species. Sample 428 was isolated from sheep which may explain why a positive result was seen with primer set L, with the sequence of the PCR product showing the *MHA_1205* gene. The NCBI searches may not have identified a *gly1* gene within *M. glucosida* due to lack of information on the database. Horizontal gene transfer may indeed have led to the presence of an exact copy of the *MHA_1205* Gly1 gene in the *M. glucosida* strain that was analysed. Alternatively it could suggest that the sample was contaminated with *M. haemolytica* due to co-isolation, although prior screening did not detect this.

Isolate 362 was identified as *M. haemolytica,* yet was negative for the *gly1* genes screened for. It does contain the paralogue and both the haemoglobin receptor HmbR and the periplasmic hemin binding protein PhuT showing that it is *M. haemolytica*. This isolate was bovine in origin, whereas 351 was isolated from an ovine host. There are reports of differences between *Mannheimia* isolated from the two host species at protein level, for example with OmpA proteins. The serotype of both isolates was unknown. The differences observed between *M. haemolytica* isolates found in different hosts could account for this variation, as could the serotype of the isolates.

Isolate 296 was contaminated with another bacterial strain. It presented as large colonies that outgrew *M. haemolytica*. PCR using primers for the *MHA_1205* gene resulted in the corresponding sequence, indicating that *M. haemolytica* was present in

the culture. Sequencing was performed with primer set D, intended to screen for the gene coding for MHA_1205. It instead amplified the gene *COI_1584*, although this has two silent mutations at the gene level compared to that sequence (Figure 6.3).

Figure 6.3 Screening Isolate 296 for MHA_1205.

Silent mutations in the D296 isolate shown in green with corresponding amino acid in capitals below. The sequence aligned with COI_1584 (Genbank accession: CP004753) rather than MHA_1205.

Primer set 'M' was one of the sets designed to detect the protein designated L280_13705, however, the sequencing showed that although there was a substitution at the gene level, this was silent and the translated nucleotide sequence exactly matched that of MHA_1205 in nine of the isolates, with identical or similar sequences for COI_1584 in a further six (Figure 6.4). This was the same for sets D, J and L in isolate 351, suggesting they were identifying the same gene.

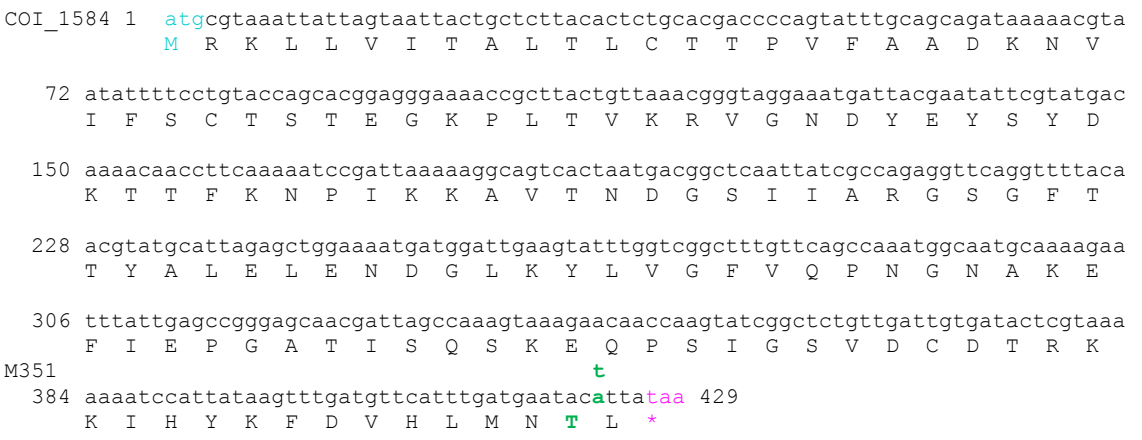


Figure 6.4. Sequence Alignment of 351 Aligned with COI_1584.

Genomic sequencing of the 351 gene amplified by primer set M aligned with the COI_1584 (top panel) showed a single substitution A-G. The substitution was shown to be silent with the base and corresponding amino acid in green. This is shown with the amino acid sequence alignment (bottom panel).

The primers designed to detect the *MHA_0579* gene did not generate any PCR products when reactions were tried under a number of conditions on the AHVLA isolates. As a positive control, the original genomic DNA from *Mannheimia haemolytica* (from which the MHA_0579 protein was cloned for this project) was used to ensure the fault was not with the primers. This produced bands of an appropriate MW on the agarose gel as expected with primer sets A, B and C (see Figure 6.5). Subsequent reactions using this set of primers on the RVC isolates were more successful, identifying identical genes in 12 out of 19 isolates.

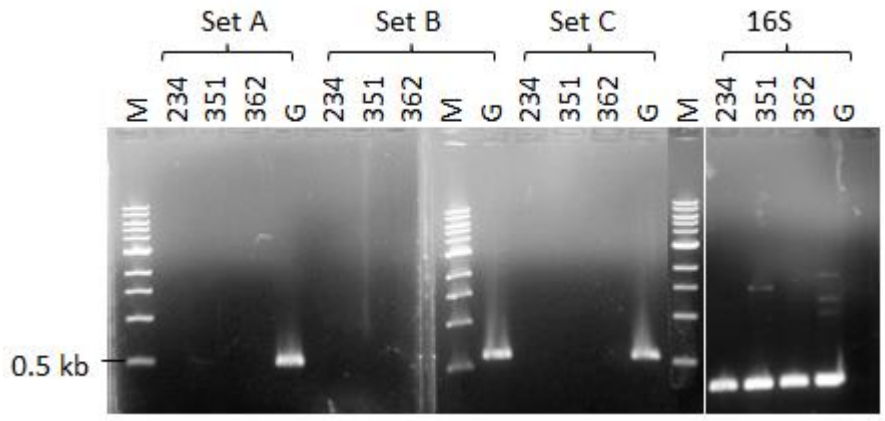


Figure 6.5. PCR Screening for MHA_0579.

Primer sets A, B and C were designed to detect the gene coding for the MHA_0579 Gly1 homologue. 'M' indicates the 1 kb marker used, 'G' indicates the original genomic DNA control and 234, 351 and 362 were the isolates screened. The 16S primers showed that the PCR was a success. No MHA_0579 gene was found in the AHVLA isolates. The primers detected the gene in the original genomic DNA, proving that they functioned. Samples separated by agarose gel electrophoresis.

One fraction of interest amplified by primer set A was observed in isolate MH4 (see Figure 6.6) with a number of substitutions at the gene level leading to a slightly different amino acid sequence. The alignment highlights the substitutions N38D, I95V, Q96K and D103S. As the structure of this protein is currently unknown, it was impossible to judge whether the substitutions were in regions of importance to its structure or function. Isolate A1 also had the D103S substitution. Isoleucine and valine are both hydrophobic, and commonly found in substitutions due to their similar properties. The other changes in properties due to the amino acid polymorphisms are alterations in charge status caused by variations in the amino acid side chains, which could conceivably alter protein structure or function.

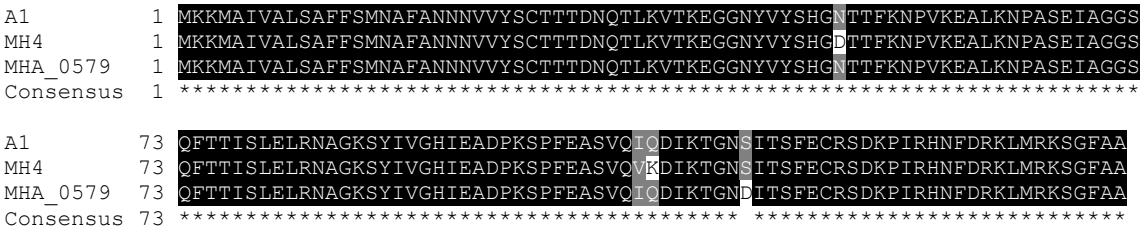


Figure 6.6 MHA_0579 Alignment with MH4 and A1 Genes.

MH4 had four amino acid polymorphisms from asparagine to aspartic acid N50D, isoleucine to valine I107V, glutamine to lysine Q108K and aspartic acid to serine D115S. A1 had a single amino acid alteration from aspartic acid to serine D115S.

Primers designed to identify the genes coding for proteins MHA_1205, COI_1584 and L280_13705 (Sets D, L and M) showed a lot of cross-over. *MHA_1205* was detected by the primers in 10 of the 25 isolates with primer set D as anticipated (Figure 6.7). The gene was also detected by Set L (in 9 of 25) and Set M (in 9 of 25). The majority of sequences found with sets D and M correspond to one another, with the exception of Set D detecting MHA_1205 in A1, where the other primers found COI_1584. Set M also detected COI_1584 in strain A9 and MH8 where Set D did not. Set L detected COI_1584 in 9 of the isolates and MHA_1205 in another 9, although these differed slightly from those found with the other two primer sets. Primer sets did not detect any genes corresponding to L280_13705. See Table 6.3 for a summary of these results.

	Primer set		
	D	L	M
A12	COI_1584*	COI_1584	MHA_1205*
A1	MHA_1205	COI_1584	COI_1584
A9		COI_1584	COI_1584
CO0600	MHA_1205	COI_1584	MHA_1205*
CO4178	MHA_1205	MHA_1205	MHA_1205
CO3983	MHA_1205	MHA_1205	MHA_1205
Gilly	MHA_1205	MHA_1205	MHA_1205
CO0353	MHA_1205	MHA_1205	MHA_1205
MH4 A8	COI_1584	COI_1584	COI_1584*
MH8 A5		COI_1584	COI_1584
A6	MHA_1205	MHA_1205	MHA_1205
CO2008	MHA_1205	MHA_1205	MHA_1205
CO1250	MHA_1205	MHA_1205	MHA_1205
234	COI_1584*	COI_1584**	COI_1584***
235		COI_1584	
296	COI_1584*		
351	COI_1584	COI_1584*	COI_1584
362		MHA_1205	
428		MHA_1205	

Table 6.3. Summary of PCR Screening with Primer Sets D, L and M.

This table summarises the sequencing results of the PCR products of primer sets D, L and M. Identical genes are coloured in green (MHA_1205) and blue (COI_1584). Genes with polymorphisms are identified with *. In 234, each gene had a different sequence hence multiple *, **, ***. Those left blank did not detect a gene.

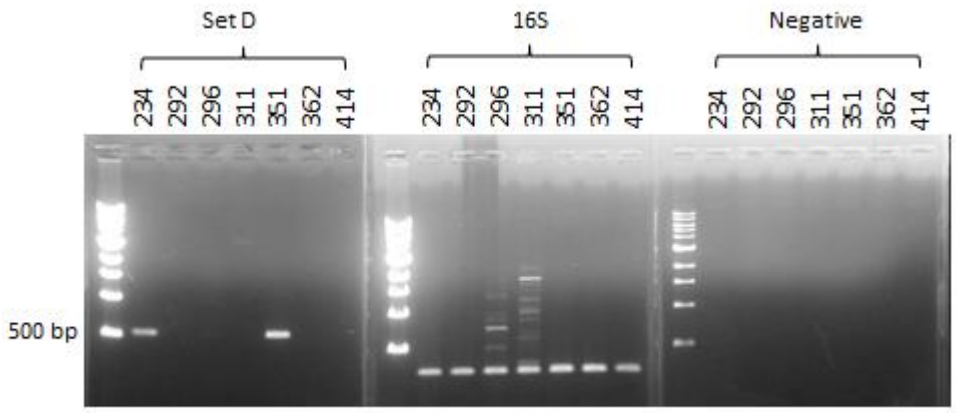


Figure 6.7. MHA_1205 Isolate Screening.

Primer set D was used to screen the isolate genomes for the gene coding for protein MHA_1205. The PCR amplified a section in isolates 234 and 351 initially believed to be the *MHA_1205* gene. Subsequent screening showed that in isolate 351 this gene was actually that coding for COI_1584. The negative control here was the identical PCR mix to that used for Set D minus DNA.

Primer set M detected a similar gene to COI_1584 seen in isolate 234 although a number of substitutions were found. The majority of these were within the predicted signal peptide up to the predicted cleavage site with the exception of isoleucine- valine substitutions. As mentioned previously, this substitution is common as these amino acids share similar properties (Figure 6.8).

```
M234         1 MRKLLGITVLTLCTTSAFATDKNVVFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKKAVTNDGSIIARG
COI_1584     1 MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKKAVTNDGSIIARG
Consensus    1 ***** ** ******  ** **** ********************************************

M234        73 SGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQSKEQPSIGSVDCDTRKKIHYKFDVHLMNTL
COI_1584    73 SGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQSKEQPSIGSVDCDTRKKIHYKFDVHLMNTL
Consensus   73 ************************************************************************
```

Figure 6.8. Sequencing of 234 with Primer Set M.

The translated nucleotide sequence generated by primer set M was different to that of the expected protein. The first 4 amino acid substitutions were within the signal peptide region (up to AFA or VFA respectively). The A20T substitution was within the processed protein but situated just after the predicted cleavage site. A further substitution I25V was also present. Predicted signal peptide underlined in red.

Primer Set J was designed to screen the isolates for COI_1584 as found in the A2 ovine strain on the NCBI database, which appeared to differ in the flanking regions when compared to other sequences available in the database. The expected product was a 631 bp band as visualised on an agarose gel. However, what was seen was actually two DNA bands of approximately 650 bp and 2.5 kbp. The lower band was extracted from the gel, the DNA purified and sent for sequencing. The sequence of the lower MW band was poor for the majority of the isolates; in 234 and 351 the sequence was sufficiently clear to allow comparison with known genes (Figure 6.9). The gene detected in isolate 234 did not match any sequences from the database, however, it was identical to COI_1584 from position 26 onwards. This was the same sequence identified with primer Set M (Figure 6.8), so it was possible that primer set J was amplifying the same gene.

```
234           1 MRKLLGITVLTLCTTSAFATDKNVVFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKK
COI_1584      1 MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKK
MHA1205       1 MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKK
mh0603        1 MRKLLVITALTLCTTPVFAADKNVIFSCTSTEGKPLTVKRVGNDYEYSYDKTTFKNPIKK
L280_13705    1 MRKLLGITVLTLCTTSAFATDKNVVFSCTSTEGKPLTVKRVGNDYEYSYDKTIFKNPIKK
consensus     1 *****.**.******..**.****.**********************************

234          61 AVTNDGSIIARGSGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQSKEQPSIG
COI_1584     61 AVTNDGSIIARGSGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQSKEQPSIG
MHA1205      61 AVTNDGSIIARGSGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQSKEQPSIG
mh0603       61 AVTNDGSIIARGSGFTTYALELENDGLKYLVGFVQPNGNAKEFIEPGATISQRKEQPSIG
L280_13705   61 AVTNDGSIIARGSGFTTYALELKNDGLKYVVGFVQPNGNSKEFIEPGATISRDNDQPSIG
consensus    61 **************************************************.*******

234         121 SVDCDTRKKIHYKFDVHLMNTL
COI_1584    121 SVDCDTRKKIHYKFDVHLMNTL
MHA1205     121 SVDCDTRKKSHYKFDVHLMNTL
mh0603      121 SVDCDTRKKSHYKFDVHLMNTL
L280_13705  121 SVDCDPYKKLYL----------
consensus   121 *********.************
```

Figure 6.9. Sequencing of 234 with Primer set J.

This primer set was designed to detect the gene coding for protein COI_1584. In all positive reactions, two bands were seen, one at approximately 2.5 kbp and one at the anticipated size of approximately 650 bp. Sequencing was attempted on other genomes but was unsuccessful.

Identical or very similar genes to the MHpara sequence identified in the initial screening were detected in 17 of the 25 isolates. Of the 8 that did not contain MHpara genes, only 4 were *M. haemolytica*. Two of the MHpara genes that were detected had different amino acid sequences (Figure 6.10). The isoleucine-phenylalanine change observed in the database genome sequences was only detected in one of the 15 MHpara genes that were identical at all other points. This shows that in the samples collected the isoleucine variant was the most common.

As mentioned in Chapter 3 Section 3.2.1, a potential homologue of MHpara was discovered with a 'frame shift' caused by a GA deletion at the gene level resulting in a shorter N terminal region but a near-identical sequence from the proposed cleavage site of the signal peptide on the longer protein. The shorter protein was not predicted to have a signal peptide at the N terminal, which would prevent the protein from being directed to the correct location within the cell. As errors can sometimes occur on the database it was not known if this variation in sequence was truly found in nature, yet sequencing of isolate 362 uncovered this same phenomenon. The putative RBS for this gene was discussed in Chapter 3. The genes on the database were flanked by the genes coding for HmbR and PhuT, suggesting it was in the same locus as the full-length version. It is unknown whether this shorter protein would be functional, and the predicted tertiary structure using the Phyre2 server does not cover the full length of the amino acid sequence to allow for comparison.

```
H362        1--------------------MQQVNKVMQEAHESTVKTSTVKYSCQNGKKLSVKYGFNEQGIPTYAEAKLSGK
MHpara      1 MKKLMIFATTAMIVSNLAHAAGEQSDAREAHESTVKTSTVKYSCQNGKKLSVKYGFNKQGIPTYAEAKLSGK
CO0600      1 MKKLMIFATTAMIVSNLAHAAGEQSDAREAHESTVKTSTVKYSCQNGKKLSVKYGFNKQGIPTYAEAKLSGK
consensus   1 ......................*****************************.**************

H362       73 KRFMPINLYTTDATGTNFGDENNFSLYGDPMEFTNHRKASVNIQSPASEILYKGCTPQK
MHpara     73 KRFMPINLYTTDATGTNFGDENNFSLYGDPMEFTNHRKASVNIQSPASEILYKGCTPQK
CO0600     73 KRFMPFNLYTTDATGTNFGDENNFSLYGDPMEFTNHRKASVNIQSPASEILYKGCTPQK
consensus  73 ************************************************************
```

Figure 6.10. MHpara Alignment with 362 and CO0600 Translated Genes.

ORF finder was used to translate the genomic sequence where GA was missing from the sequenced PCR product at position 84-85 along with an A-G substitution at position 171. These account for the initial change in the N terminus and the lysine to glutamic acid difference at amino acid level. No signal peptide was predicted with the 362 MHpara, and the non-matching segment was not found in the NCBI database. This was the same as CO2392. Primer set H was used.

Further to the above sequencing, a brief investigation into the genomic location of the *MHpara* gene compared to the *hmbr* gene of the isolates was performed. As seen in Chapter 3 in the flanking gene analysis (Figure 3.5) it was noted that one of the interesting genes found in the *M. haemolytica* genome, the *MHpara* gene, was adjacent to a haemoglobin receptor gene, *hmbr*.

To ascertain if this genomic arrangement was repeated in the isolates screened, PCR was performed using the forward primer for *MHpara*, initially used to clone this gene, and the forward primer for *hmbr*, again used for cloning. If the two genes were located in a similar position to those seen in the database, the product size would equal approximately 2.7 kbp consisting of each gene and a 210 bp linker region. As can be seen in Figure 6.11, the product band created by PCR using the HmbR primers alone was a lower molecular weight (approx. 2.1 kbp) than the product band created by using both the MHpara and HmbR forward primers (approx. 2.7 kbp). This suggested that in the isolates that contain both genes they were situated close to each other and in the expected orientation as described in the database genome sequences. Confirmation of the presence of both *MHpara* and *hmbr* genes was performed by partial sequencing of the PCR products with the primers used for the initial reaction.

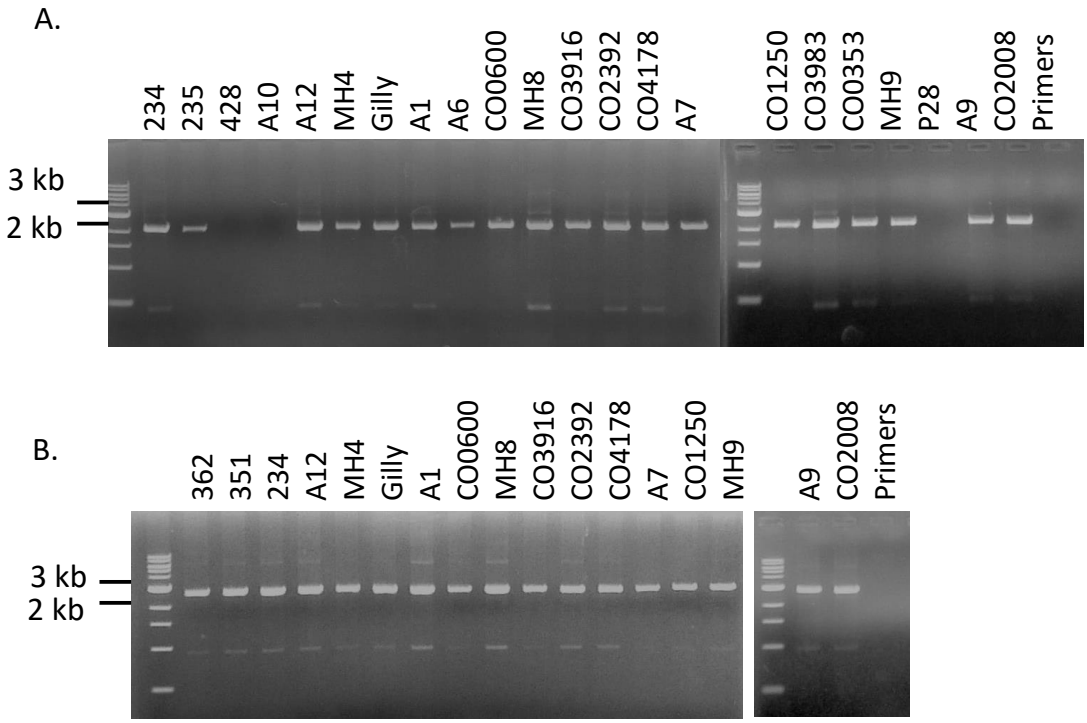


Figure 6.11. MHpara and HmbR PCR Results.

Panel **A.** shows the PCR products using the HmbR primers only as analysed by agarose gel electrophoresis. Panel **B.** shows the products created with both HmbR and MHpara primers. Those not found to have MHpara in initial sequencing were not included in panel B. The 'Primers' lane had no DNA in the reaction mix.

The possible Gly1 proteins discovered in the database searches were putative homologues of the neisserial Gly1 proteins. Genomic screening of the donated *Mannheimia haemolytica* isolates revealed that the majority of them contained at least two of the potential Gly1 homologues. Exceptions were isolates found to be other bacterial species and isolates A7, CO2392, CO3916, 362 and MH9 A2.

This trend was supported by the database findings, which overall placed two or three Gly1 proteins along with the MHpara protein in each genome sequence with one exception. These homologues were not within the same genetic locus and were separated by between approximately 8 kbp and 500 kbp based on the database numbering system. The database A2 isolates had either the *MHpara* gene alone or with one COI_1584 gene. No A7 genomic sequence has been deposited in the NCBI database with which to compare the isolate screening results, and the serotypes of CO2392 and

CO3916 were not known. If the genes identified using primer sets D and L were in fact the same gene, the results supported the database genome findings (Figure 6.12).

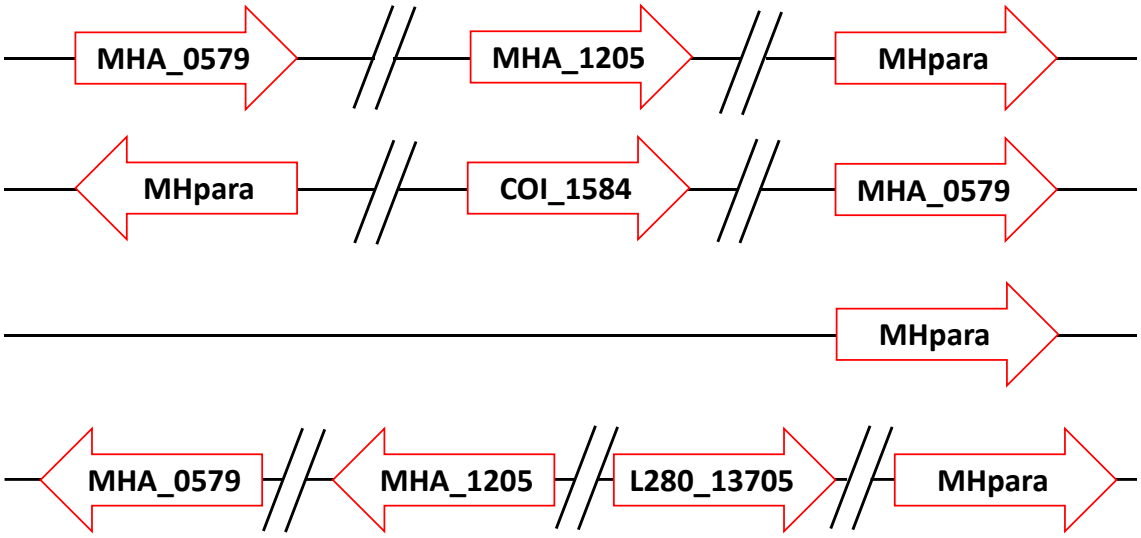


Figure 6.12. A Summary of the Presence of Genes of Interest in Available Genome Sequences.

Gly1 was not present in every sequence on the NCBI database. All but one *M. haemolytica* isolate contained MHpara in some form. Fifteen isolates contained MHA_0579, fourteen MHA_1205, three COI_1584, and three L280_13705. The genes of interest were not found within similar genomic locations to each other as indicated by the broken lines.

6.2.3 Protein Expression

In order to further investigate the presence of these genes, the *M. haemolytica* isolates were grown in a rich medium (BHI) overnight at optimum temperature. The liquid cultures were normalised by measuring optical density and 1 ml of each was centrifuged down to pellet the bacteria. The cell pellets were lysed and the proteins separated by SDS-PAGE. Western blots were performed using rat anti-MHA_1205 C-His whole antisera, and purified rabbit anti-MHpara and rabbit anti-MHA_0579 C-His to detect any Gly1 protein expression.

MHpara was detected in the Gilly strain, which was isolated due to its enhanced pathogenicity. It was also detected in the A1 and A6 serotyped strains which again are known to be the major isolates involved in bovine respiratory disease (Figure 6.13). The strains of *M. haemolytica* that were of unknown serotype in which the antibodies detected proteins of interest were CO3983 in two out of three repeats, and CO2392, CO4178 in one of three repeats. In both MH8 (serotype A5) and MH9 (serotype A2) MHpara was also detected in one of three repeats.

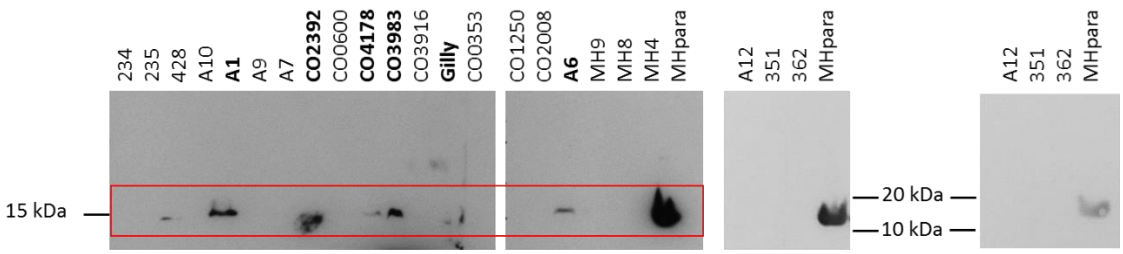


Figure 6.13. Western Blot on Whole Cell Lysate of Isolates to Detect MHpara.

Proteins were separated by SDS-PAGE prior to analysis by western blot. MHpara was detected in A6 and Gilly in each repeat. It was detected in A1 and CO3983 in two out of the three repeats. Cells were grown in iron replete conditions. The panel to the right showed a repeat of the western blot but with the anti-MHpara pre-incubated with MHpara protein. N = 3.

Pre-incubation of the antibodies with purified MHpara or MHA_0579 C-His proteins reduced or eliminated the detection of the protein of interest on the western blots. This was also seen with the cell lysates that were positive for MHpara, supporting the idea that the proteins detected were indeed MHpara. A number of strains that were found to contain the gene for MHpara after PCR screening did not show positive results for the protein. Nor was any MHA_0579 protein detected by the specific antibodies in any isolate (Figure 6.14). PCR screening had shown that MHA_0579 was present in a minimum of thirteen of the isolates. As we have shown some evidence suggesting MHA_0579 interacts with hemin in Chapter 4, it could be suggested that the expression is down-regulated under iron replete conditions. This is a common occurrence with proteins involved in haem uptake and storage, which are often tightly regulated. Additionally, mh0603 was observed to be upregulated in conditions of iron paucity (Roehrig et al., 2007).

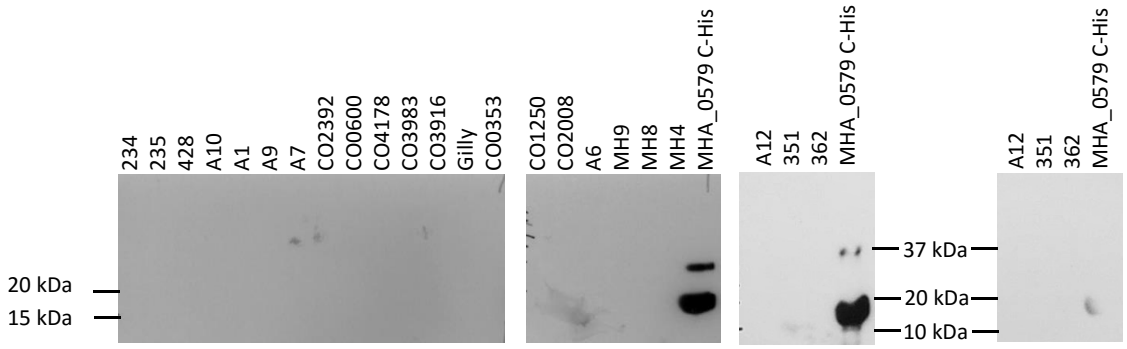


Figure 6.14. MHA_0579 Expression in Isolates Under Iron Replete Conditions.

Western blot detection was carried out on proteins separated by SDS-PAGE. No MHA_0579 was detected in any of the lysed samples. N = 3.

MHA_1205 protein was detected using whole rat serum. This led to a fair amount of non-specific binding. The antisera was pre-incubated with *E. coli* whole cell lysate, which decreased the additional protein bands (Figure 6.15, panels B and 3). This suggested that the remaining bands present on the western blot were MHA_1205 or COI_1584. The PCR screenng did not detect either of these or any similar proteins in isolates A10, A7, CO2392, CO3916 or MH9. MHA_1205 was detected in isolate 235 in the PCR screening

however no protein band was seen in the western blot. The size of the band was a lower MW in the isolate lysate than the pure MHA_1205 C-His protein. This difference could be attributed to the linker region and polyhistidine tag, which is an additional 14 amino acids, with an approximate MW of 1.8 kDa.

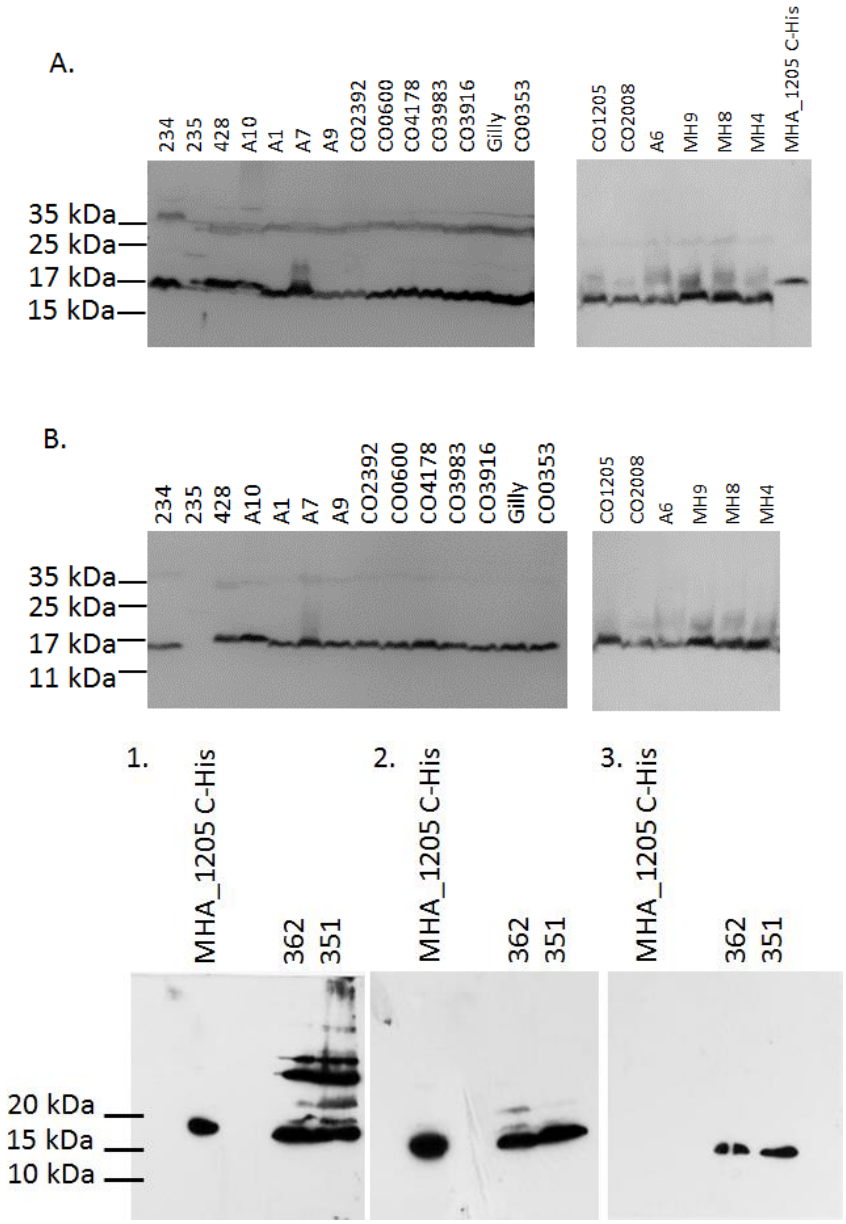


Figure 6.15 MHA_1205 and COI_1584 Expression in Isolates Under Iron Replete Conditions.
Antisera against MHA_1205 C-His was used to identify expression in the *M. haemolytica* isolate whole cell lysate of MHA_1205 and COI_1584 High levels of background (**Panel A. and 1.**) were reduced by pre-incubating the antisera with *E. coli* cell lysate along with MHA_1205 C-His (**Panel B. and 2.**). Antisera incubated with both *E. coli* cell lysate and MHA_1205 C-His (**Panel 3.**). The MW of the proteins detected was marginally lower than MHA_1205 C-His.

6.3 Discussion

At one stage, *Mannheimia* species such as *M. haemolytica*, *M. varigena* and *M. glucosida* were all classified as serotypes of the biovar 'A' of *Pasteurella haemolytica*. Reclassification and analysis based on phenotypic and 16S typing showed that *M. haemolytica* and *M. glucosida* are the most similar of the *Mannheimia* species, although their 16S sequences do differ slightly (Angen et al., 1999). While other *Mannheimia* species are also found as commensals in cattle, sheep and, in the case of *M. varigena*, pigs, *M. varigena* is associated with disease such as pneumonia and septicaemia and *M. glucosida* with ovine mastitis.

It is interesting to note that *M. varigena* (235) and *M. glucosida* (428) were found to contain COI_1584 or MHA_1205 genes, respectively. There were four assemblies of *M. varigena* available on the NCBI database, and a BLASTp search uncovered potential Gly1 homologues in isolates 1261 (accession number AHG73145.1) and 1388 (accession number AHG79471.1). These protein sequences were identical to those found in numerous other bacteria including *P. multocida* and *B. trehalosi*, which could suggest an annotation error. The genes identified share 88% identity with COI_1584 and from this information it was not clear if these sequences were reliable evidence of the presence of a Gly1 gene in these species. The COI_1584-like protein found in isolate 235 was identical to the *Mannheimia* protein. Contamination of the sample is possible, as the leukotoxin sequence was identical to that found in *M. haemolytica*, which was not the sequence found in *M. varigena* on the database. The database Lkt sequences between *Mannheimia* species share 94% identity overall but just 83% at the short region of 66 amino acids sequenced with the Lkt primers used in this study, which supports the idea that the sample was either contaminated or had been mis-labelled at isolation. This was further reinforced by the fact that bacterial growth was seen on MacConkey agar which would not be expected with *M. varigena*, although the morphology of the colonies besides this is reported to be very similar between species.

The *M. glucosida* leukotoxin sequence was very similar to that of *M. haemolytica,* with 97% identity (see Figure 6.16). The section of gene used for screening with the 'Lkt' primers is 100% identical to the *M. haemolytica* sequence, however further screening

over a more diverse section of the gene named 'Lkt2' was able to differentiate between these two species. This found that the sample was indeed *M. glucosida*. The lack of whole genome sequences in the NCBI database makes it difficult to know whether this proposed *gly1* gene is commonly found in this species or if this is an unusual event. *M. haemolytica* is naturally competent as described by Gioia *et al* based on the discovery of genes involved in DNA uptake (Gioia et al., 2006). The fact that *M. glucosida* is believed to be so similar to *M. haemolytica* based on other sequencing results could indicate genetic element transfer in the opposite direction from *M. haemolytica*, and provide an explanation for the identical gene for MHA_1205.

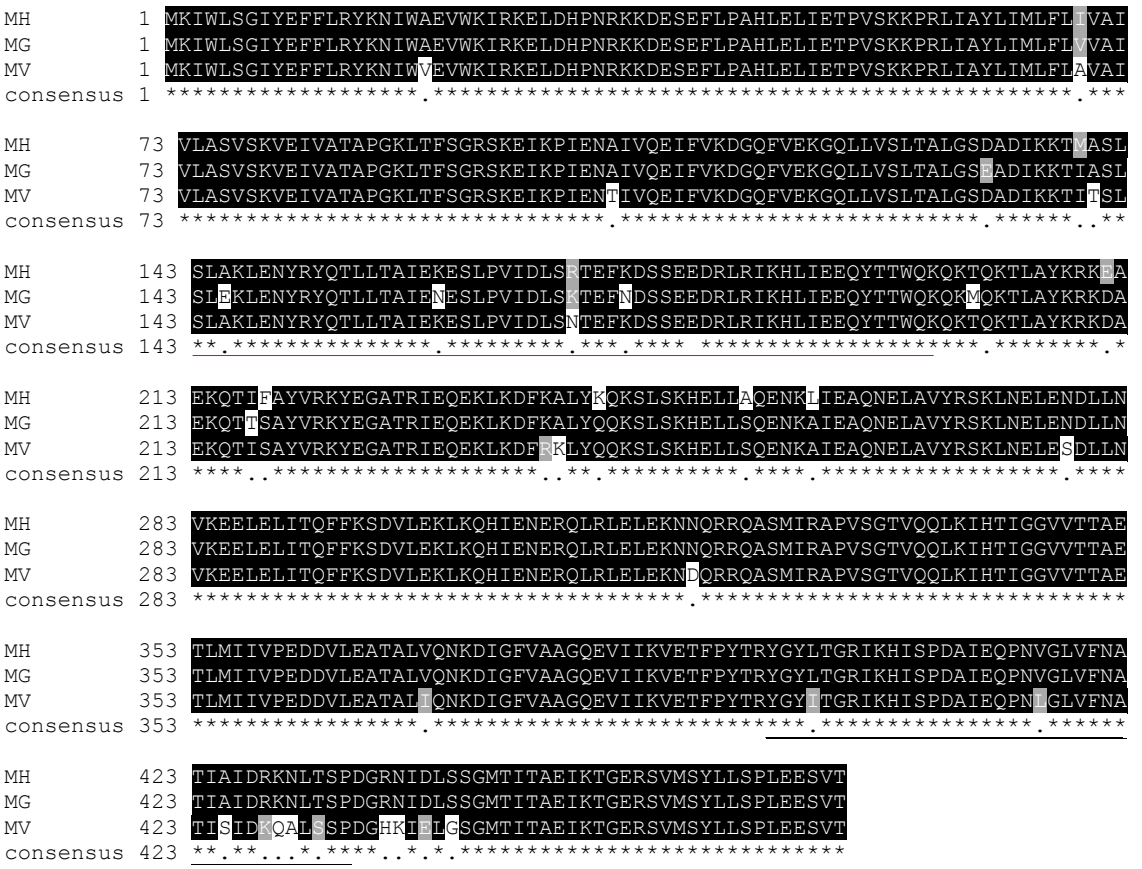


Figure 6.16. Alignment of Leukotoxin or 'Hemolysin D' from *M. haemolytica* (MH), *M. varigena* (MV) and *M. glucosida* (MG).

The sequence identity of the leukotoxin proteins (also known as hemolysin D) in mannheimia species. MH Lkt accession AAL12786.1, MV accession WP_025235106.1, MG accession AAL12798.1. The section underlined in red was Lkt2 for MG screening and in black for MH screening.

Isolate 235 did not contain a MHpara-like protein sequence but both 235 and 428 contained PhuT and HmbR. This supports the idea that 235 was actually *M. haemolytica*, although there is insufficient information available on the NCBI database to know whether *M. glucosida* contains these two proteins or similar homologues.

An interesting observation was that the phenylalanine form of MHpara was only observed in one of the MHpara isolates screened, with the isoleucine form predominant. The larger amino acid, phenylalanine, was predicted to partially occlude the pore of the MHpara protein tertiary structure (Chapter 3, Figure 3.24). This could indicate that the larger amino acid is not the common biological one and supports the hypothesis that it is a rare single nucleotide polymorphism from an A to a T and may lead to a malfunctioning protein. The protein used throughout the project for structure-function experiments was the isoleucine form. It could be interesting to produce and label the phenylalanine form and analyse interactions of this protein compared to the isoleucine version.

The strain labelled 'Gilly' was a particularly virulent strain that caused a bad pneumonia outbreak on a farm (verbal communication from Prof. Rycroft, RVC). While the serotype of this strain was unknown, it can be assumed that it was A1, A2 or A6 as these are the most commonly occurring isolates involved in pathogenesis and disease as discussed in Chapter 1 section 1.2.2.1.

On the NCBI database, the strain isolated from deer was a novel isolate. It does not appear to have even distant Gly1 homologues or paralogues, and only bears 86% identity to the coding sequences of other strains of *M. haemolytica* (Lawrence et al., 2014). This could explain the lack of Gly1 homologues identified in this particular strain.

Results of the PCR isolate screening with different primer sets detected a gene more than once as multiple primer sets were used to account for up/ down stream variability in genome sequence. While it is possible that the gene was duplicated within the genome, the primer sequences could be detecting the exact same gene in the DNA. The primer sets A, B and C, (MHA_0579) and D, L and M (MHA_1205 and similar) were designed to cover the differences observed in the flanking regions of the database genomes available at the time, and it is possible these detected the same gene. Gene

duplication is an important mechanism in the adaptation and evolution of species, with examples observed in the bovine commensal *Mannheimia succiniciproducens* (Hong et al., 2004).

While the majority of the potential duplicate genes were identical to each other, five of the isolates had different genetic sequences that when translated gave rise to identical protein sequences. Subsequent genome additions to the NCBI database showed that these primer sets identify similar regions of the genome on the database available. This explains the 'identical' genes discovered as being the same gene, but not the lack of the gene in some or differences observed between others. If the genes were duplicated, the duplication event possibly included the regions up- and down-stream of the coding region, as the primers were situated either around 50 bp or 100 bp up- or down-stream from the start and stop codons. As illustrated in Figure 6.12, the genomes available on the database have multiple copies of Gly1 homologues. It was initially hypothesised that the MHA_1205, COI-1584, mh0603 and L280 proteins were actually the same protein that had undergone minimal mutations. The presence of a number of versions of these similar genes in the screening results is in line with the database information, as half of the isolates showed different sequences. Further investigations would clarify whether the remaining 9 isolates had multiple copies of the gene within their genomes of if it was in fact the same gene being detected multiple times.

Primer set J was designed to detect COI_1584 based on the A2 ovine strain on the NCBI database, which had slightly different flanking regions of DNA. These primer sets amplified multiple bands in the majority of isolates, and the only isolates that sequences were able to be deduced for were two ovine strains- 234 and 351. This may be a result of the differences observed between ovine and bovine strains (Hounsome et al., 2011).

The detection of MHpara in all *M. haemolytica* strains except A10, A6, CO3983, CO0353, 235 and 428 using PCR, but not in all western blots of the cell lysates was interesting; even more so because the antibodies used in the western blot detected MHpara in A6 and CO3983. This could be due to differences in the DNA sequences flanking the genes, as the primers designed were specific to the sequences in the NCBI database, and

changes in the DNA could alter the binding of these, preventing the reaction from occurring.

The PCR screening protocol was informative however, due to the nature of primer specification, it is liable to give false negatives when screening genomic sequences for the presence of genes. Mismatches may occur within the base pairing region. Southern blotting and FISH hybridisation methods can probe for the sequences and do not rely on primer specification for regions of DNA surrounding the gene.

The cultures were grown in the same incubator in 37° C with 5% CO_2 on a shaker in 25 ml universal containers. BHI medium was used throughout. The cultures grew at different rates, and although efforts were made to adjust the density of the cell culture to approximately A_{600} 0.5 for each experiment, it is possible that there was some variation in cell density. The cells were harvested at stationary phase after around 16 hours of growth. Alternatively, the differing growth rates could be related to the serotype of the bacterial isolate with some expressing these possible adhesins and others not. As discussed in Chapter 1, Section 1.2.2.2, there are differences between the serotypes, even between those found in different hosts, and especially in the pathogenicity of each. A known virulence factor is cellular adhesion, and if MHpara is indeed an adhesin there could be some correlation between the MHpara gene and pathogenicity. This would be very interesting as a potential target for new BRD therapies.

Screening of the field isolates of *M. haemolytica* revealed that the majority (19 of 25) contained one or more of the highly similar Gly1 homologues. This suggests that these proteins play an important role in *M. haemolytica* biology and potentially pathogenesis. They could therefore present possible therapeutic targets if antibodies raised against the proteins confer complement mediated killing. Further investigations into the prevalence and expression of these proteins in the common pathogenic strains is required to understand the full potential of these putative Gly1 proteins.

In conclusion, the screening of these isolates has shown that the Gly1 homologues or paralogue could have potential in terms of new targeted therapies to prevent or tackle the main bacterial cause of BRD.

7. Conclusion

The basis of this project was to investigate the putative Gly1 homologues uncovered via NCBI BLAST database searches of available genomes. No similar sequences to the bacterial Gly1 proteins have been uncovered to date in eukaryotic organisms, enhancing the potential of this proposed family of proteins as novel specific bacterial targets to prevent disease in the mammalian hosts. Neisserial proteins have previously been shown to bind to haem and erythrocytes, and antibodies against these demonstrated complement mediated killing of the bacteria. Our hypothesis was that putative Gly1 proteins found in other animal pathogens shared similar cellular functions, such as haem binding, and could therefore play an important role in the biology of a wide range of Gram negative organisms.

7.1 Identification of Gly1 Homologues and Structural Investigations

Database analysis uncovered proteins of similar sequence in many Gram-negative bacteria, all with unknown function. Initial studies therefore centred on analysis of key pathogens and the structural similarity of their proteins. This showed that the proteins shared a secondary structure composed of majority beta strands. In addition to this, MHA_0579 and MHA_1205 of *M. haemolytica* and SARI of *S. enterica arizonae IIIa* readily formed dimers, as was previously observed with the neisserial Gly1 proteins. Higher oligomers were also detected in the concentrated protein stocks. APP Gly1 and MHpara did not oligomerise, the first indication that these two proteins differed from the rest of those under investigation. The multimers were formed primarily by disulphide bonds. It was proposed that the dimers observed in *N. meningitidis* Gly1 were the result of domain swapping of the two beta sheets forming the tertiary structure, although this was not observed in the crystal structure solved by a collaboration at The University of Sheffield.

Disulphide bonds generally stabilise the protein, and the ability of these proteins to form multimers in this manner may give an indication of the method of binding the ligand i.e. Gly1 in *N. meningitidis* interacted in a 2:1 ratio with hemin (Sathyamurthy, 2011). The oligomerisation occurring via disulphide bonds can also help provide evidence for the

predicted cellular location of the protein. The bacterial cytoplasm is a reducing environment, therefore it is not common for proteins to form bonds which can, as demonstrated with the reducing agent DTT, be readily broken down in this fashion.

To further analyse the similarities between these proteins and build on the information provided by tertiary structure prediction software, optimisation of the X-ray crystallography attempted would be beneficial. The project to deduce tertiary structures of the proposed Gly1 homologues investigated in this thesis is ongoing. The MHpara protein tertiary structure was solved and found to resemble the anti-parallel beta barrel seen with both *N. meningitidis* and *N. gonorrhoeae* Gly1 proteins. The central pore of the MHpara protein was smaller at less than 10 Å in diameter and with loops appearing to partially occlude it in the form crystallised. These protein loops may be mobile and involved in the function of MHpara. As the holo forms of proteins are in some cases more stable, continuing with the haem soaked crystals could allow diffraction to occur to a sufficient degree for the structure to prove solvable. It would be of interest to both compare to known Gly1 structures and to database proteins. The cellular location of the Gly1 homologues was predicted as OM for MHA_0579, the neisserial Gly1s and SARI, however, MHA_1205 and COI_1584 were predicted to be periplasmic. This again suggests that these proteins have differing cellular functions.

7.2 Studying the Function of Gly1-like Proteins

The initial studies by Arvidson *et al* showed that certain isolates of *N. gonorrhoeae* were associated with haemolytic activity not commonly observed with this species, one of which was found to contain Gly1. This provided another indication that the protein may be involved in haem acquisition, although subsequent haemolytic activity observed in the *E. coli* cells overexpressing this protein may have been an artefact of the overexpression itself.

In *M. haemolytica*, as the name suggests, β-hemolysis is a method used by the bacteria to acquire haem from host erythrocytes. The leukotoxin produced by *M. haemolytica* is known to be the causative factor of this weak haemolysis (Murphy et al., 1995). Bacteria can produce multiple haemolytic toxins; α, β, γ and δ hemolysis has been described, opening this avenue as a potential direction of investigation.

7.2.1 Haem Interactions

This project found that MHA_0579 interacted with haem or hemin in the majority of assays conducted, with MHA_1205 and SARI proteins demonstrating interaction with haem or hemin in one fewer experiment. These results were suggestive of the proteins functioning as haem binding proteins in the bacterial haem acquisition system.

Investigations into protein expression in the *M. haemolytica* isolates found that while the MHA_0579 protein was not detectable under iron replete conditions, MHA_1205 (or one of the very similar homologues) was detected in each of the isolates under the same growth conditions. It would be interesting to assess this on a transcriptional level to quantify up- or down- regulation of these and other *gly1* genes under differing growth conditions. mh0603, one of the proteins similar to MHA_1205 was observed to be upregulated under conditions of iron paucity, suggesting a role in iron acquisition (Roehrig et al., 2007). The original level of transcription was not stated, so it is possible that there was a low level of expression under iron rich conditions as observed with the positive results of the western blots detecting MHA_1205 in the isolated analysed in this thesis.

While the results of the absorbance spectra and peroxidase activity assays suggested haem binding, any positive results observed in the hemin agarose pull down were abolished upon cleavage of the polyhistidine tag. Unfortunately there was insufficient material to perform enterokinase cleavage on a large scale, and attempts to overexpress and purify MHA_1205 in its wild type form were unsuccessful due to poor expression and low solubility. MHA_1205 C-His did not show a shift in the absorbance spectra suggesting that the polyhistidine tag was not interfering with the results of this assay, and only a minor shift was recorded in the hemin spectra when incubated with SARI C-His. This suggests that while the interactions observed in the hemin agarose assay resulted from the presence of the polyhistidine tag on the protein, this was not the cause of positive results for other assays. Aggregation has been observed in MHA_1205 C-His under certain conditions (Sayers lab, unpublished), and it is possible that the solid surface of the agarose beads encourages this. Alternatively, the immobilised haem may have prevented natural binding, whereas this steric hindrance was not an issue in the

other assays performed. MHpara and APP Gly1 proteins, neither of which were cloned with a polyhistidine tag, were not observed to interact with haem or hemin in any experiments conducted. MHA_0579 is involved in bacterial haem uptake, and we can hypothesise that it is down-regulated during growth in iron rich environments. It is unlikely that it is a storage protein based on this, yet it is possible that MHA_1205 or COI_1584 proteins are as the western blots show these in all isolates.

Due to the diversity of haem binding sites and mechanisms it is very hard to analyse the interactions occurring based on predicted structures, and even to predict interactions in the first place. Solving the tertiary structure of these potential haem binding proteins would enable a more in-depth investigation into the binding properties of Gly1 proteins that look to interact with haem. However, even then the binding clefts can vary slightly from protein to protein, so attempts would need to be made to solve the structure of the holo-protein to discover the exact means of binding.

Ideally we would have liked to create Gly1 knockouts in the isolates that contain said genes in order to analyse growth on different media, however, this was not possible due to the time constraints of getting ethical approval for genetic modification once the haem interactions had been reported. It would be interesting to investigate the knock out versions for both the haem binding proteins and MHpara and cellular adhesion (see Chapter 5).

7.2.2 *M. haemolytica* Paralogue Function

MHpara did not interact with haem, however it did show structural similarity to a lysozyme inhibitor from *E. coli,* PliC. The MHpara-lysozyme interaction assay was inconclusive, showing a higher percentage of the lane total formed by the higher MW band, although this may have resulted from variability between samples. Time constraints precluded probing with lysozyme antibodies at the time of investigation therefore it was not clear whether the higher MW bands were composed of both proteins, as could be deduced by western blot. It may be that the presence of lysozyme encouraged the formation of MHpara multimers, which were not observed in the dimerisation experiments in Chapter 5.

Lysozyme inhibitors can be periplasmic or membrane lipoproteins, and while the cellular location remains unknown, MHpara does not contain the lipoprotein motif in the signal peptide as there is an absence of cysteines (Hayashi and Wu, 1990). Subsequent investigations showed that MHpara in fact acts as an adhesin. This potential function was investigated due to the identification of the *N. meningitidis* putative Gly1 paralogue 'ACP' as an adhesin, which shares some identity with the MHpara protein as discussed in Chapter 5. In the PliC paper the flanking genes of both PliC and MliC were identified as those involved in bacterial macrophage survival, which would function in conjunction with the lysozyme inhibitor to defeat the host defence systems (Callewaert et al., 2008).

MHpara, as mentioned in Chapter 3, was located upstream of an apparent haem uptake locus of *M. haemolytica* and downstream of a collection of transporter proteins. Based on this evidence it is more likely that the MHpara protein is an adhesin than a lysozyme inhibitor. As discussed in Chapter 1, adhesins are known virulence factors for bacteria. If expressed when the bacteria are attempting to colonise the host the MHpara could present an attractive target. In addition to this proposed important function, genome screening showed that genes coding for MHpara were present in the majority of isolates, and, while this could allow for an effective vaccine, western blot analysis did not detect protein expression in all isolates with the gene. If it is periplasmic like PliC then it is unlikely to be a real adhesin. It is worth considering multiple functions which, although rare, are increasingly being described. This may be in the form of proteins with multiple ligands (pleiotropism) or moonlighting proteins as discussed in a review by Huberts *et al* (Huberts and van der Klei, 2010). It is plausible that the protein could both be involved in adhesion and, as a result of this placing of the bacterial cell in a region containing lysozyme, also protecting the bacteria from this particular host defence mechanism.

7.2.3 Predicting Protein Function Related to Cellular Location

When considering the potential function of each protein, it is interesting to refer back to the flanking genes annotated on the NCBI database. These showed hypothetical membrane proteins down stream of both MHA_1205 and MHA_0579. It is possible that these as yet uncharacterised proteins are involved in the bacterial haem uptake system. As mentioned, MHA_0579 is predicted to be an outer membrane protein, however, as

a small protein it is unlikely that MHA_0579 is an outer membrane receptor in its own right. It could therefore interact with the larger putative membrane protein located adjacent in the genome to bring haem into the cell. SignalP could not identify transmembrane domain motifs on any Gly1 proteins. These membrane proteins do not have any similarity to currently described haem or haemoglobin receptors, but as further haem uptake mechanisms with a variety of different protein sequences and structures are being discovered this does not discount them as being involved in haem acquisition. MHA_1205 is predicted to be a periplasmic protein. Again, MHA_1205 is smaller than the haem PBP's characterised from Gram-negative bacteria to date. It was observed to interact with PhuT in the BLItz experiment, which could suggest it acts as a periplasmic storage protein or chaperone. Optimisation of haemoglobin receptor expression in the competent *E. coli* cells may shed some light on Gly1 activity in the haem uptake system.

For further work to corroborate the results from the software in Table 3.4, a potential step is separation of the periplasm from the outer and inner membranes of the cell of isolates found to express the proteins of interest (Chapter 6). This would enable identification of the cellular location of these proteins. Alternatively, labelled antibodies could be used with fluorescence microscopy on fixed bacterial cells.

7.3 The Potential of Gly1 Homologues as Vaccine Targets

Studies into *N. meningitidis* Gly1 highlighted its potential as a target for antimicrobial activity. It is believed to be an outer membrane protein, giving it potential as a surface antigen of the bacteria and allowing for easier detection by antibodies. It is understood that the capsule of *M. haemolytica* can mask some of these surface antigens. There is evidence that this capsule can interfere with complement mediated bactericidal activity, as demonstrated by capsular/ acapsular mutant comparisons, which would potentially reduce the efficacy of this method (McKerral and Lo, 2002). *N. meningitidis*, however, also boasts a capsule amongst its defence repertoire, which could interfere with the bactericidal effect of the membrane attack complex formed via the complement cascade. Investigations in the Sayers lab showed that neisserial Gly1 antibodies induced complement mediated bacterial killing, reinforcing the notion that the protein is

situated on the OM and also providing evidence that the capsule does not prevent this process from occurring (Weirzbicka, 2014).

The viability as a vaccine candidate would again depend upon the expression levels of these potential target proteins. As observed in the western blots, no MHA_0579 was detected although MHA_1205 or similar appeared to be throughout, and MHpara was detected inconsistently in a number of isolates. It would be revealing to analyse the serotypes of the isolates to identify whether these proteins are expressed in the more pathogenic isolates; A1, A2 and A6. Additionally, quantification of expression in different growth conditions could further investigate the necessity of the individual proteins in cellular survival. It is possible that the proteins already induce antibody production during host infection and it would therefore be interesting to analyse serum of naïve, infected and convalescent animals. OmpA is highly immunogenic and is the major known protein of *M. haemolytica* that antibodies are raised against. Trials of OmpA as a vaccine candidate showed immune responses in both mice and cattle along with complement mediated killing when incubated with serum from the vaccinated animals (Ayalew et al., 2011).

The putative Gly1 homologue from *S. enterica arizonae IIIa*, SARI, was also found to bind haem and is therefore hypothesised to be involved in haem acquisition. While this was initially thought to have less potential as a basis for a vaccine as the proportion of disease caused by this bacteria is very low, the antisera cross reactivity showed that rat anti-SARI antisera was able to detect the wild type varieties of MHA_1205, MHA_0579 and even slightly *N. meningitidis* Gly1. This antigen therefore presents a potential broad-range vaccine component. Further studies investigating the immunogenicity and bactericidal activity would be required to corroborate this evidence.

7.4 Summary

This project has shown that the putative periplasmic haem binding protein PhuT does in fact bind hemin, and has a predicted structure similar to that of other PBP's. Secondary structure prediction suggested a majority alpha helical shape, however, spectral analysis showed more beta than alpha even though the CD spectrum had a shape characteristic of alpha helical protein. This was possibly due to the reference proteins within the

database on the Dichroweb analysis software. Periplasmic haem binding proteins have been observed to bind the Fe^{2+} and Fe^{3+} oxidation states of haem with different affinities, with many binding hemin preferentially. PhuT caused a shift in the hemin spectrum but not the haem spectrum. It is therefore possible that the *M. haemolytica* PhuT protein has a higher affinity for hemin than haem. Confirmation of the cellular location would provide the final information required to state that this protein is a periplasmic haem binding protein of *M. haemolytica*. The predicted haem PBP, PhuT has been shown to bind haem supporting the annotation on the NCBI database.

APP Gly1 function is still unknown. No haem binding was observed in any studies performed, and no further ideas were gleaned from the results of this project. Preliminary investigations into APP Gly1- erythrocyte binding showed a slight shift in positive population in flow cytometry but this was not conclusive as a population shift was also observed with some controls (data not shown). Further preliminary adhesion studies using mouse anti-APP Gly1 antisera again did not appear to show any interaction with the porcine trachea, although the staining had a high level of background (data not shown). Unfortunately the low amount of antisera available limited the possibility of experimental repeats post-optimisation. The putative Gly1 homologue from *A. pleuropneumoniae* has, however, been discounted as a Gly1 family protein, and this investigation has cast doubt over whether it even exists within this bacterial genome, but *A. suis* or *A. equuis.*

Overall this thesis has described previously unknown basic functions of five bacterial proteins and added to the structural knowledge of a putative family of haem binding proteins 'Gly1' found in Gram-negative bacteria. It has highlighted the potential of these proteins as vaccine candidates based on the discovery of their genes in bacterial isolates.

(USDA), U. S. D. O. A. 2013. Feedlot 2011: Part I Management Practices on Feedlots with a Capacity of 1000 or More Head [Online]. http://www.aphis.usda.gov/animal_health/nahms/feedlot/downloads/feedlot2011/Feed11_dr_PartI.pdf. [Accessed August 2015].

ABDUL-GADER, A., MILES, A. J. & WALLACE, B. A. 2011. A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics*. 12, 1630-6.

ACKERMAN, DERSCHEID & ROTH, 2011. Innate immunology of bovine respiratory disease. *Vet Clin North Am Food Anim Pract*, 26, 215-228.

ADAMCZAK, R., POROLLO, A. & MELLER, J. 2005. Combining prediction of secondary structure and solvent accessibility in proteins. *Proteins*, 59, 467-75.

ADLAM, C., KNIGHTS, J. M., MUGRIDGE, A., LINDON, J. C., BAKER, P. R., BEESLEY, J. E., SPACEY, B., CRAIG, G. R. & NAGY, L. K. 1984. Purification, characterization and immunological properties of the serotype-specific capsular polysaccharide of *Pasteurella haemolytica* (serotype A1) organisms. *J Gen Microbiol*, 130, 2415-26.

ADLAM, C., KNIGHTS, J. M., MUGRIDGE, A., LINDON, J. C., WILLIAMS, J. M. & BEESLEY, J. E. 1986. Purification, characterization and immunological properties of the serotype-specific capsular polysaccharide of *Pasteurella haemolytica* serotype A7 organisms. *J Gen Microbiol,* 132, 1079-87.

ADUSU, T.E., CONLON, P.D., SHEWEN, P.E. & BLACK, WD, 1994. *Pasteurella haemolytica* leukotoxin induces histamine release from bovine pulmonary mast cells. Can J Vet Res. 58, 1-5.

AJITO, T., HAGA, Y., HOMMA, S., GORYO, M. & OKADA, K. 1995. Immunohistological evaluation on lung lesions of pigs intranasally inoculated with *Actinobacillus pleuropneumoniae* type 1. *J. Vet. Med. Sci*, 58, 297- 303.

ALEXANDER, T. W., COOK, S. R., YANKE, L. J., BOOKER, C. W., MORLEY, P. S., READ, R. R., GOW, S. P. & MCALLISTER, T. A. 2008. A multiplex polymerase chain reaction assay for the identification of *Mannheimia haemolytica*, *Mannheimia glucosida* and *Mannheimia ruminalis*. Vet Microbiol, 130, 165-75.

ALLEN, J. W., LEACH, N. & FERGUSON, S. J. 2005. The histidine of the c-type cytochrome CXXCH haem-binding motif is essential for haem attachment by the *Escherichia coli* cytochrome c maturation (Ccm) apparatus. *Biochem J*. 389 (2), 587-92.

ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25, 3389-402.

ANDREWS, S. M. & POLLARD, A. J. 2014. A vaccine against serogroup B *Neisseria meningitidis*: dealing with uncertainty. *Lancet Infect Dis*, 14, 426-34.

ANGEN, O., MUTTERS, R., CAUGANT, D. A., OLSEN, J. E. & BISGAARD, M. 1999. Taxonomic relationships of the *[Pasteurella] haemolytica* complex as evaluated by DNA-DNA hybridizations and 16S rRNA sequencing with proposal of *Mannheimia haemolytica* gen. nov., comb. nov., *Mannheimia granulomatis* comb. nov., *Mannheimia glucosida* sp. nov., *Mannheimia ruminalis* sp. nov. and *Mannheimia varigena* sp. nov. *Int J Syst Bacteriol*, 49 Pt 1, 67-86.

ANZALDI, L. L. & SKAAR, E. P. 2010. Overcoming the heme paradox: heme toxicity and tolerance in bacterial pathogens. *Infect Immun*. 78 (12), 4977-89.

ARVIDSON, C. G., KIRKPATRICK, R., WITKAMP, M. T., LARSON, J. A., SCHIPPER, C. A., WALDBESER, L. S., O'GAORA, P., COOPER, M. & SO, M. 1999. *Neisseria gonorrhoeae* mutants altered in toxicity to human fallopian tubes and molecular characterization of the genetic locus involved. *Infect Immun*, 67, 643-52.

ARVIDSON, D. N., PEARSON, R. F. & ARVIDSON, C. G. 2003. Purification, characterization and preliminary X-ray crystallographic studies on *Neisseria gonorrhoeae* Gly1ORF1. *Acta Crystallogr D Biol Crystallogr*. 59 (4), 747-8.

AUGER, E., DESLANDES, V., RAMJEET, M., CONTRERAS, I., NASH, J. H., HAREL, J., GOTTSCHALK, M., OLIVIER, M. & JACQUES, M. 2009. Host-pathogen interactions of *Actinobacillus pleuropneumoniae* with porcine lung and tracheal epithelial cells. *Infect Immun*, 77, 1426-41.

AYALEW, S., SHRESTHA, B., MONTELONGO, M., WILSON, A. E. & CONFER, A. W. 2011. Immunogenicity of *Mannheimia haemolytica* recombinant outer membrane proteins serotype 1-specific antigen, OmpA, OmpP2, and OmpD15. Clin Vaccine Immunol, 18, 2067-74.

BALL, H. J., CONNOLLY, M. & CASSIDY, J. 1993. *Pasteurella haemolytica* serotypes isolated in Northern Ireland during 1989-1991. *Br Vet J*, 149, 561-70.

BANDARA, A. B., LAWRENCE, M. L., VEIT, H. P. & INZANA, T. J. 2003. Association of *Actinobacillus pleuropneumoniae* capsular polysaccharide with virulence in pigs. *Infect Immun*, 71, 3320-8.

BHATIA, B., MITTAL, K. R. & FREY, J. 1991. Factors involved in immunity against *Actinobacillus pleuropneumoniae* in mice. *Vet Microbiol*, 29, 147-58.

BISGAARD, M. 1993. Ecology and significance of pasteurellaceae in animals. *Zentralblatt für Bakteriologie*, 279, 7-26.

BRIONES, V., TELLEZ, S., GOYACHE, J., BALLESTEROS, C., DEL PILAR LANZAROT, M., DOMINGUEZ, L. & FERNANDEZ-GARAYZABAL, J. F. 2004. Salmonella diversity associated with wild reptiles and amphibians in Spain. *Environ Microbiol*, 6, 868-71.

CAILLET-SAGUY, C., PICCIOLI, M., TURANO, P., LUKAT-RODGERS, G., WOLFF, N., RODGERS, K. R., IZADI-PRUNEYRE, N., DELEPIERRE, M. & LECROISEY, A. 2012. Role of the iron axial ligands of heme carrier HasA in heme uptake and release. *J Biol Chem*. 287 (32), 26932-43.

CALLEWAERT, L., AERTSEN, A., DECKERS, D., VANOIRBEEK, K. G., VANDERKELEN, L., VAN HERREWEGHE, J. M., MASSCHALCK, B., NAKIMBUGWE, D., ROBBEN, J. & MICHIELS, C. W. 2008. A new family of lysozyme inhibitors contributing to lysozyme tolerance in gram-negative bacteria. *PLoS Pathog*, 4, e1000019.

CAR, B. D., SUYEMOTO, M. M., NEILSEN, N. R. & SLAUSON, D. O. 1991. The role of leukocytes in the pathogenesis of fibrin deposition in bovine acute lung injury. *Am J Pathol*, 138, 1191-8.

CDC. 2015. Meningococcal Disease: Technical & Clinical Information [Online]. Centers for Disease Control and Prevention: US Government. [Accessed 19 July 2015 2015].

CHAE, C. H., GENTRY, M. J., CONFER, A. W. & ANDERSON, G. A. 1990. Resistance to host immune defense mechanisms afforded by capsular material of *Pasteurella haemolytica*, serotype 1. *Vet Microbiol*, 25, 241-51.

CHAN, K., BAKER, S., KIM, C. C., DETWEILER, C. S., DOUGAN, G. & FALKOW, S. 2003. Genomic comparison of *Salmonella enterica* serovars and *Salmonella bongori* by use of an *S. enterica* serovar typhimurium DNA microarray. *J Bacteriol*, 185, 553-63.

CHIANG, C. H., HUANG, W. F., HUANG, L. P., LIN, S. F. & YANG, W. J. 2009. Immunogenicity and protective efficacy of ApxIA and ApxIIA DNA vaccine against *Actinobacillus pleuropneumoniae* lethal challenge in murine model. *Vaccine*, 27, 4565-70.

CHOO, K. H., TAN, T. W. & RANGANATHAN, S. 2009. A comprehensive assessment of N-terminal signal peptides prediction methods. *BMC Bioinformatics*. **10**(Suppl 15):S2.

CHU, B. C., OTTEN, R., KREWULAK, K. D., MULDER, F. A. & VOGEL, H. J. 2014. The solution structure, binding properties, and dynamics of the bacterial siderophore-binding protein FepB. J Biol Chem. 289 (42), 29219-34.

CONFER, A. W., AYALEW, S., PANCIERA, R. J., MONTELONGO, M. & WRAY, J. H. 2006. Recombinant *Mannheimia haemolytica* serotype 1 outer membrane protein PlpE enhances commercial *M. haemolytica* vaccine-induced resistance against serotype 6 challenge. *Vaccine*, 24, 2248-55.

CZJZEK, M., LETOFFE, S., WANDERSMAN, C., DELEPIERRE, M., LECROISEY, A. & IZADI-PRUNEYRE, N. 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 (4), 1176-86.

D'AMELIO, N., BONVIN, A. M., CZISCH, M., BARKER, P. & KAPTEIN, R. 2002. The C terminus of apocytochrome b562 undergoes fast motions and slow exchange among ordered conformations resembling the folded state. *Biochemistry*. 41 (17), 5505-14.

DAVIES, R. L. & BAILLIE, S. 2003. Cytotoxic activity of *Mannheimia haemolytica* strains in relation to diversity of the leukotoxin structural gene lktA. *Vet Microbiol*, 92, 263-79.

DAVIES, R. L. & DONACHIE, W. 1996. Intra-specific diversity and host specificity within *Pasteurella haemolytica* based on variation of capsular polysaccharide, lipopolysaccharide and outer-membrane proteins. *Microbiology*, 142 (Pt 7), 1895-907.

DE AMORIM, G. C., PROCHNICKA-CHALUFOUR, A., DELEPELAIRE, P., LEFEVRE, J., SIMENEL, C., WANDERSMAN, C., DELEPIERRE, M. & IZADI-PRUNEYRE, N. 2013. The structure of HasB reveals a new class of TonB protein fold. PLoS One. 8 (3), e58964.

DELANY, I., RAPPUOLI, R. & SCARLATO, V. 2004. Fur functions as an activator and as a repressor of putative virulence genes in *Neisseria meningitidis*. *Mol Microbiol*. 52 (4), 1081-90.

DENIAU, C., GILLI, R., IZADI-PRUNEYRE, N., LETOFFE, S., DELEPIERRE, M., WANDERSMAN, C., BRIAND, C. & LECROISEY, A. 2003. Thermodynamics of heme binding to the HasA(SM) hemophore: effect of mutations at three key residues for heme uptake. *Biochemistry*, 42, 10627-33.

DOUSSE, F., THOMANN, A., BRODARD, I., KORCZAK, B. M., SCHLATTER, Y., KUHNERT, P., MISEREZ, R. & FREY, J. 2008. Routine phenotypic identification of bacterial species of the family Pasteurellaceae isolated from animals. *J Vet Diagn Invest*, 20, 716-24.

EAKANUNKUL, S., LUKAT-RODGERS, G. S., SUMITHRAN, S., GHOSH, A., RODGERS, K. R., DAWSON, J. H. & WILKS, A. 2005. Characterization of the periplasmic heme-binding protein shut from the heme uptake system of *Shigella dysenteriae*. *Biochemistry*, 44, 13179-91.

ENZ, S., BRAND, H., ORELLANA, C., MAHREN, S. & BRAUN, V. 2003. Sites of interaction between the FecA and FecR signal transduction proteins of ferric citrate transport in *Escherichia coli* K-12. *J Bacteriol*, 185, 3745-52.

EUZEBY, J. P. 1997. List of Bacterial Names with Standing in Nomenclature: a folder available on the Internet. *Int J Syst Bacteriol*, 47, 590-2.

EVANS, N. J., HARRISON, O. B., CLOW, K., DERRICK, J. P., FEAVERS, I. M. & MAIDEN, M. C. 2010. Variation and molecular evolution of HmbR, the *Neisseria meningitidis* haemoglobin receptor. *Microbiology*, 156, 1384-93.

FABER, R., HARTWIG, N., BUSBY, D. & BREDAHL, R. 1999. The Costs and Predictive Factors of Bovine Respiratory Disease in Standardized Steer Tests. Iowa State University: Beef Research Report IA. A. S. Leaflet R1648.

FLORES JIMENEZ, R. H. & CAFISO, D. S. 2012. The N-terminal domain of a TonB-dependent transporter undergoes a reversible stepwise denaturation. *Biochemistry*, 51, 3642-50.

FODOR, L., VARGA, J., HAJTOS, I. & MOLNAR, T. 1999. Serotypes of *Pasteurella haemolytica* and *Pasteurella trehalosi* isolated from farm animals in Hungary. *Zentralbl Veterinarmed B*, 46, 241-7.

FONTECAVE, M. & PIERRE, J. L. 1993. Iron: metabolism, toxicity and therapy. *Biochimie*, 75, 767-73.

FUSCO, W. G., AFONINA, G., NEPLUEV, I., CHOLON, D. M., CHOUDHARY, N., ROUTH, P. A., ALMOND, G. W., ORNDORFF, P. E., STAATS, H., HOBBS, M. M., LEDUC, I. & ELKINS, C. 2010. Immunization with the *Haemophilus ducreyi* hemoglobin receptor HgbA with adjuvant monophosphoryl lipid A protects swine from a homologous but not a heterologous challenge. *Infect Immun*, 78, 3763-72.

FUSCO, W. G., CHOUDHARY, N. R., COUNCIL, S. E., COLLINS, E. J. & LEDUC, I. 2013. Mutational analysis of hemoglobin binding and heme utilization by a bacterial hemoglobin receptor. *J Bacteriol*. 78 (9), 3763-72.

GATTO, N. T., CONFER, A. W., ESTES, D. M., WHITWORTH, L. C. & MURPHY, G. L. 2006. Lung lesions in SCID-bo and SCID-bg mice after intratracheal inoculation with wild-type or leucotoxin-deficient mutant strains of *Mannheimia haemolytica* serotype 1. *J Comp Pathol*, 134, 355-65.

GIOIA, J., QIN, X., JIANG, H., CLINKENBEARD, K., LO, R., LIU, Y., FOX, G. E., YERRAPRAGADA, S., MCLEOD, M. P., MCNEILL, T. Z., HEMPHILL, L., SODERGREN, E., WANG, Q., MUZNY, D. M., HOMSI, F. J., WEINSTOCK, G. M. & HIGHLANDER, S. K. 2006. The genome sequence of *Mannheimia haemolytica* A1: insights into virulence, natural competence, and Pasteurellaceae phylogeny. *J Bacteriol*. United States.

GREENFIELD, N. J. 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*. 188 (20), 7257-66.

GRESOCK, M. G., SAVENKOVA, M. I., LARSEN, R. A., OLLIS, A. A. & POSTLE, K. 2011. Death of the TonB Shuttle Hypothesis. *Front Microbiol*, 2, 206.

GREY, C. L. & THOMSON, R. G. 1971. *Pasteurella haemolytica* in the tracheal air of calves. *Can J Comp Med*, 35, 121-8.

GRIFANTINI, R., SEBASTIAN, S., FRIGIMELICA, E., DRAGHI, M., BARTOLINI, E., MUZZI, A., RAPPUOLI, R., GRANDI, G. & GENCO, C. A. 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, 9542-7.

GRYLLOS, I., SHAW, J. G., GAVIN, R., MERINO, S. & TOMAS, J. M. 2001. Role of flm locus in mesophilic Aeromonas species adherence. *Infect Immun*, 69, 65-74.

HARRISON, L. H., TROTTER, C. L. & RAMSAY, M. E. 2009. Global epidemiology of meningococcal disease. *Vaccine*, 27 Suppl 2, B51-63.

HARRIS and JANZEN. 1989. The *Haemophilus somnus* disease complex (haemophilosis): A review. *Canadian Veterinary Journal,* 30, 816-822.

HAYASHI, S. & WU, H. C. 1990. Lipoproteins in bacteria. *J Bioenerg Biomembr*, 22, 451-71.

HECKENBERG, S. G., DE GANS, J., BROUWER, M. C., WEISFELT, M., PIET, J. R., SPANJAARD, L., VAN DER ENDE, A. & VAN DE BEEK, D. 2008. Clinical features, outcome, and meningococcal

genotype in 258 adults with meningococcal meningitis: a prospective cohort study. *Medicine*. 87, 185-92.

HO, W. W., LI, H., EAKANUNKUL, S., TONG, Y., WILKS, A., GUO, M. & POULOS, T. L. 2007. Holo- and apo-bound structures of bacterial periplasmic heme-binding proteins. *J Biol Chem*, 282, 35796-802.

HOLLENSTEIN, K., DAWSON, R. J. & LOCHER, K. P. 2007. Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol*. 17 (4), 142-8.

HONG, S. H., KIM, J. S., LEE, S. Y., IN, Y. H., CHOI, S. S., RIH, J. K., KIM, C. H., JEONG, H., HUR, C. G. & KIM, J. J. 2004. The genome sequence of the capnophilic rumen bacterium *Mannheimia succiniciproducens*. *Nat Biotechnol*, 22, 1275-81.

HOUNSOME, J. D., BAILLIE, S., NOOFELI, M., RIBOLDI-TUNNICLIFFE, A., BURCHMORE, R. J., ISAACS, N. W. & DAVIES, R. L. 2011. Outer membrane protein A of bovine and ovine isolates of *Mannheimia haemolytica* is surface exposed and contains host species-specific epitopes. *Infect Immun*, 79, 4332-41.

HSUAN, S. L., KANNAN, M. S., JEYASEELAN, S., PRAKASH, Y. S., MALAZDREWICH, C., ABRAHAMSEN, M. S., SIECK, G. C. & MAHESWARAN, S. K. 1999. *Pasteurella haemolytica* leukotoxin and endotoxin induced cytokine gene expression in bovine alveolar macrophages requires NF-kappaB activation and calcium elevation. *Microb Pathog*, 26, 263-73.

HUBERTS, D. H. & VAN DER KLEI, I. J. 2010. Moonlighting proteins: an intriguing mode of multitasking. *Biochim Biophys Acta*, 1803, 520-5.

HUNG, M. C., HECKELS, J. E. & CHRISTODOULIDES, M. 2013. The adhesin complex protein (ACP) of *Neisseria meningitidis* is a new adhesin with vaccine potential. *MBio*. 4 (2), e00041-13.

INZANA, T.J., GOGOLEWSKI, R.P. & CORBEIL, L.B. 1992. Phenotypic Phase Variation in *Haemophilus somnus* Lipooligosaccharide during Bovine Pneumonia and after in vitro passage. *Infection and Immunity*. 60, 2943-2951.

IZADI-PRUNEYRE, N., HUCHE, F., LUKAT-RODGERS, G. S., LECROISEY, A., GILLI, R., RODGERS, K. R., WANDERSMAN, C. & DELEPELAIRE, P. 2006. The heme transfer from the soluble HasA hemophore to its membrane-bound receptor HasR is driven by protein-protein interaction from a high to a lower affinity binding site. *J Biol Chem*. 281 (35), 25541-50.

JACKSON, L. A., DUCEY, T. F., DAY, M. W., ZAITSHIK, J. B., ORVIS, J. & DYER, D. W. 2010. Transcriptional and functional analysis of the *Neisseria gonorrhoeae* Fur regulon. *J Bacteriol*, 192, 77-85.

JACOB, F. & MONOD, J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*, 3, 318-56.

JANA, B., MANNING, M. & POSTLE, K. 2011. Mutations in the ExbB cytoplasmic carboxy terminus prevent energy-dependent interaction between the TonB and ExbD periplasmic domains. *J Bacteriol*. 193 (20), 5649-57.

JANSEN, R., BRIAIRE, J., SMITH, H. E., DOM, P., HAESEBROUCK, F., KAMP, E. M., GIELKENS, A. L. & SMITS, M. A. 1995. Knockout mutants of *Actinobacillus pleuropneumoniae* serotype 1 that are devoid of RTX toxins do not activate or kill porcine neutrophils. *Infect Immun*, 63, 27-37.

JIAN, Z., ALLEY, M. R. & MANKTELOW, B. W. 1991. Experimental pneumonia in mice produced by combined administration of *Bordetella parapertussis* and *Pasteurella haemolytica* isolated from sheep. *J Comp Pathol*, 104, 233-43.

KATRIBE, E., BOGOMOLNAYA, L. M., WINGERT, H. & ANDREWS-POLYMENIS, H. 2009. Subspecies IIIa and IIIb Salmonellae are defective for colonization of murine models of salmonellosis compared to *Salmonella enterica* subsp. I serovar typhimurium. *J Bacteriol*, 191, 2843-50.

KATSUDA, K., KOHMOTO, M., MIKAMI, O., TAMAMURA, Y. & UCHIDA, I. 2012. Plasmid-mediated florfenicol resistance in *Mannheimia haemolytica* isolated from cattle. *Vet Microbiol*, 155, 444-7.

KELLEY, L. A., MEZULIS, S., YATES, C. M., WASS, M. N. & STERNBERG, M. J. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. England.

KHAN, A. G., SHOULDICE, S. R., KIRBY, S. D., YU, R. H., TARI, L. W. & SCHRYVERS, A. B. 2007. High-affinity binding by the periplasmic iron-binding protein from *Haemophilus influenzae* is required for acquiring iron from transferrin. *Biochem J*. 10 (6), 845-58.

KISIELA, D. I. & CZUPRYNSKI, C. J. 2009. Identification of *Mannheimia haemolytica* adhesins involved in binding to bovine bronchial epithelial cells. *Infect Immun*, 77, 446-55.

KLIMA, C. L., ALEXANDER, T. W., HENDRICK, S. & MCALLISTER, T. A. 2014a. Characterization of *Mannheimia haemolytica* isolated from feedlot cattle that were healthy or treated for bovine respiratory disease. *Can J Vet Res*, 78, 38-45.

KLIMA, C. L., ZAHEER, R., COOK, S. R., BOOKER, C. W., HENDRICK, S., ALEXANDER, T. W. & MCALLISTER, T. A. 2014b. Pathogens of bovine respiratory disease in North American feedlots conferring multidrug resistance via integrative conjugative elements. *J Clin Microbiol*, 52, 438-48.

KLITGAARD, K., FRIIS, C., ANGEN, O. & BOYE, M. 2010. Comparative profiling of the transcriptional response to iron restriction in six serotypes of *Actinobacillus pleuropneumoniae* with different virulence potential. *BMC Genomics*, 11, 698.

KOLKER, S., ITSEKZON, T., YINNON, A. M. & LACHISH, T. 2012. Osteomyelitis due to *Salmonella enterica subsp. arizonae*: the price of exotic pets. *Clin Microbiol Infect*, 18, 167-70.

KUCEROVA, Z., HRADECKA, H., NECHVATALOVA, K. & NEDBALCOVA, K. 2011. Antimicrobial susceptibility of *Actinobacillus pleuropneumoniae* isolates from clinical outbreaks of porcine respiratory diseases. *Vet Microbiol*, 150, 203-6.

LAFLEUR, R. L., MALAZDREWICH, C., JEYASEELAN, S., BLEIFIELD, E., ABRAHAMSEN, M. S. & MAHESWARAN, S. K. 2001. Lipopolysaccharide enhances cytolysis and inflammatory cytokine induction in bovine alveolar macrophages exposed to *Pasteurella (Mannheimia) haemolytica* leukotoxin. *Microb Pathog*, 30, 347-57.

LAKEW, W., GIRMA, A. & TRICHE, E. 2013. *Salmonella enterica* Serotype Arizonae Meningitis in a Neonate. *Case Rep Pediatr*, 2013, 813495.

LARSON, R. L. & STEP, D. L. 2012. Evidence-based effectiveness of vaccination against *Mannheimia haemolytica*, *Pasteurella multocida*, and *Histophilus somni* in feedlot cattle for mitigating the incidence and effect of bovine respiratory disease complex. *Vet Clin North Am Food Anim Pract*, 28, 97-106, 106e1-7, ix.

LAWRENCE, P. K., BEY, R. F., WIENER, B., KITTICHOTIRAT, W. & BUMGARNER, R. E. 2014. Genome Sequence of a Presumptive *Mannheimia haemolytica* Strain with an A1/A6-Cross-Reactive Serotype from a White-Tailed Deer (Odocoileus virginianus). *Genome Announc*, 27;2(2). e00114-14.

LEE, I. & DAVIES, R. L. 2011. Evidence for a common gene pool and frequent recombinational exchange of the tbpBA operon in *Mannheimia haemolytica*, *Mannheimia glucosida* and *Bibersteinia trehalosi*. *Microbiology*, 157, 123-35.

LEE, K. S., RAYMOND, L. D., SCHOEN, B., RAYMOND, G. J., KETT, L., MOORE, R. A., JOHNSON, L. M., TAUBNER, L., SPEARE, J. O., ONWUBIKO, H. A., BARON, G. S., CAUGHEY, W. S. & CAUGHEY, B. 2007. Hemin interactions and alterations of the subcellular localization of prion protein. *J Biol Chem*. 292 (50), 36525-33.

LEITE, F., O'BRIEN, S., SYLTE, M. J., PAGE, T., ATAPATTU, D. & CZUPRYNSKI, C. J. 2002. Inflammatory cytokines enhance the interaction of *Mannheimia haemolytica* leukotoxin with bovine peripheral blood neutrophils in vitro. *Infect Immun*, 70, 4336-43.

LETOFFE, S., HEUCK, G., DELEPELAIRE, P., LANGE, N. & WANDERSMAN, C. 2009. Bacteria capture iron from heme by keeping tetrapyrrol skeleton intact. *Proc Natl Acad Sci U S A*. 106 (28), 11719-24.

LETOFFE, S., NATO, F., GOLDBERG, M. E. & WANDERSMAN, C. 1999. Interactions of HasA, a bacterial haemophore, with haemoglobin and with its outer membrane receptor HasR. *Mol Microbiol*, 33, 546-55.

LEWIS, L. A., ROHDE, K., GIPSON, M., BEHRENS, B., GRAY, E., TOTH, S. I., ROE, B. A. & DYER, D. W. 1998. Identification and molecular analysis of lbpBA, which encodes the two-component meningococcal lactoferrin receptor. *Infect Immun*, 66, 3017-23.

LEYSEN, S., VAN HERREWEGHE, J. M., CALLEWAERT, L., HEIRBAUT, M., BUNTINX, P., MICHIELS, C. W. & STRELKOV, S. V. 2011. Molecular basis of bacterial defense against host lysozymes: X-ray structures of periplasmic lysozyme inhibitors PliI and PliC. *J Mol Biol*. 405 (5), 1233-45.

LI, T., BONKOVSKY, H. L. & GUO, J. T. 2011. Structural analysis of heme proteins: implications for design and prediction. *BMC Struct Biol*, 11, 13.

LIN, M. H., CHANG, Y. C., HSIAO, C. D., HUANG, S. H., WANG, M. S., KO, Y. C., YANG, C. W. & SUN, Y. J. 2013. LipL41, a hemin binding protein from *Leptospira santarosai* serovar Shermani. *PLoS One*. 8 (12), e83246.

LIU, X., OLCZAK, T., GUO, H. C., DIXON, D. W. & GENCO, C. A. 2006. Identification of amino acid residues involved in heme binding and hemoprotein utilization in the *Porphyromonas gingivalis* heme receptor HmuR. Infect Immun. 74 (2), 1222-32.

LO, R. Y. & SORENSEN, L. S. 2007. The outer membrane protein OmpA of *Mannheimia haemolytica* A1 is involved in the binding of fibronectin. *FEMS Microbiol Lett*, 274, 226-31.

LOCHER, K. P., REES, B., KOEBNIK, R., MITSCHLER, A., MOULINIER, L., ROSENBUSCH, J. P. & MORAS, D. 1998. Transmembrane signalling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell*, 95, 771-8.

LOSINGER, W. C. 2005. Economic impacts of reduced pork production associated with the diagnosis of *Actinobacillus pleuropneumoniae* on grower/finisher swine operations in the United States. *Prev Vet Med*. 68 (2-4), 181-93.

MANDRELL, R. E., MCLAUGHLIN, R., ABA KWAIK, Y., LESSE, A., YAMASAKI, R., GIBSON, B., SPINOLA, S. M. & APICELLA, M. A. 1992. Lipooligosaccharides (LOS) of some *Haemophilus* species mimic human glycosphingolipids, and some LOS are sialylated. *Infect Immun*, 60, 1322-8.

MASON, K. M., RAFFEL, F. K., RAY, W. C. & BAKALETZ, L. O. 2011. Heme utilization by nontypeable *Haemophilus influenzae* is essential and dependent on Sap transporter function. J Bacteriol. 193 (10), 2527-35.

MATTLE, D., ZELTINA, A., WOO, J. S., GOETZ, B. A. & LOCHER, K. P. 2010. Two stacked heme molecules in the binding pocket of the periplasmic heme-binding protein HmuT from *Yersinia pestis. J Mol Biol*, 404, 220-31.

MCCLUSKEY, J., GIBBS, H. A. & DAVIES, R. L. 1994. Variation in outer-membrane protein and lipopolysaccharide profiles of *Pasteurella haemolytica* isolates of serotypes A1 and A2 obtained from pneumonic and healthy cattle. *Microbiology*, 140 (Pt 4), 807-14.

MCKERRAL, L. J. & LO, R. Y. 2002. Construction and characterization of an acapsular mutant of *Mannheimia haemolytica* A1. *Infect Immun*, 70, 2622-9.

MEADOWS, C. 2005. Characterisation of the Emi protein and the alpha peptides of *Neisseria meningitidis*. PhD, The University of Sheffield.

MERMIN, J., HUTWAGNER, L., VUGIA, D., SHALLOW, S., DAILY, P., BENDER, J., KOEHLER, J., MARCUS, R. & ANGULO, F. J. 2004. Reptiles, amphibians, and human Salmonella infection: a population-based, case-control study. *Clin Infect Dis*, 38 Suppl 3, S253-61.

MILLIKEN, G. A. 2013. *Mannheimia haemolytica* efficacy studies demonstrating the absence of excessive interference of Titanium products with the *Mannheimia haemolytica - Pasteurella multocida* bacterin-toxoid. Elanco: Eli Lilly.

MOKRY, D. Z., NADIA-ALBETE, A., JOHNSON, M. K., LUKAT-RODGERS, G. S., RODGERS, K. R. & LANZILOTTA, W. N. 2014. Spectroscopic evidence for a 5-coordinate oxygenic ligated high spin ferric heme moiety in the *Neisseria meningitidis* hemoglobin binding receptor*. Biochim Biophys Acta*, 1840, 3058-66.

MORTON, D. J., SEALE, T. W., BAKALETZ, L. O., JURCISEK, J. A., SMITH, A., VANWAGONER, T. M., WHITBY, P. W. & STULL, T. L. 2009. The heme-binding protein (HbpA) of *Haemophilus influenzae* as a virulence determinant. *Int J Med Microbiol*, 299, 479-88.

MURPHY, G. L., WHITWORTH, L. C., CLINKENBEARD, K. D. & CLINKENBEARD, P. A. 1995. Hemolytic activity of the *Pasteurella haemolytica* leukotoxin. *Infect Immun*, 63, 3209-12.

NAKAI, K. & HORTON, P. 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci*, 24, 34-6.

ODENDAAL, M. W. & HENTON, M. M. 1995. The distribution of *Pasteurella haemolytica* serotypes among cattle, sheep, and goats in South Africa and their association with diseases. Onderstepoort *J Vet Res*, 62, 223-6.

PENG, D., MA, L. H., SMITH, K. M., ZHANG, X., SATO, M. & LA MAR, G. N. 2012. Role of propionates in substrate binding to heme oxygenase from *Neisseria meningitidis*: a nuclear magnetic resonance study. *Biochemistry*, 51, 7054-63.

PETERSEN, T. N., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 8 (10), 785-6.

POWELL, N. B., BISHOP, K., PALMER, H. M., ALA'ALDEEN, D. A., GORRINGE, A. R. & BORRIELLO, S. P. 1998. Differential binding of apo and holo human transferrin to meningococci and co-localisation of the transferrin-binding proteins (TbpA and TbpB). *J Med Microbiol*, 47, 257-64.

PUENTE-POLLEDO, L., REGLERO, A., GONZALEZ-CLEMENTE, C., RODRIGUEZ-APARICIO, L. B. & FERRERO, M. A. 1998. Biochemical conditions for the production of polysialic acid by *Pasteurella haemolytica* A2. *Glycoconj J,* 15, 855-61.

QI, Z., HAMZA, I. & O'BRIAN, M. R. 1999. Heme is an effector molecule for iron-dependent degradation of the bacterial iron response regulator (Irr) protein. *Proc Natl Acad Sci U S A*, 96, 13056-61.

RAMJEET, M., COX, A. D., HANCOCK, M. A., MOUREZ, M., LABRIE, J., GOTTSCHALK, M. & JACQUES, M. 2008. Mutation in the LPS outer core biosynthesis gene, galU, affects LPS interaction with the RTX toxins ApxI and ApxII and cytolytic activity of *Actinobacillus pleuropneumoniae* serotype 1. *Mol Microbiol*, 70, 221-35.

RAS. 2009. Acetone Precipitation of Proteins [Online]. Available: http://www.piercenet.com/files/TR0049-Acetone-precipitation.pdf 2011].

RASBAND, W. S. 1997-2012. ImageJ. U.S. National Institute of Health, Bethesda, Maryland, USA.

RIVERA-RIVAS, J. J., KISIELA, D. & CZUPRYNSKI, C. J. 2009. Bovine herpesvirus type 1 infection of bovine bronchial epithelial cells increases neutrophil adhesion and activation. *Vet Immunol Immunopathol*, 131, 167-76.

ROEHRIG, S. C., TRAN, H. Q., SPEHR, V., GUNKEL, N., SELZER, P. M. & ULLRICH, H. J. 2007. The response of *Mannheimia haemolytica* to iron limitation: implications for the acquisition of iron in the bovine lung. *Vet Microbiol*. 121 (3-4), 316-29.

SATHYAMURTHY, S. 2011. The role of Gly1, a secreted neisserial protein in meningococcal pathogenesis. PhD, The University of Sheffield.

SAYERS, J. R. & ECKSTEIN, F. 1991. A single-strand specific endonuclease activity copurifies with overexpressed T5 D15 exonuclease. *Nucleic Acids Res*, 19, 4127-32.

SCALLAN, E., HOEKSTRA, R. M., ANGULO, F. J., TAUXE, R. V., WIDDOWSON, M. A., ROY, S. L., JONES, J. L. & GRIFFIN, P. M. 2011. Foodborne illness acquired in the United States--major pathogens. *Emerg Infect Dis*, 17, 7-15.

SEO, K. W., KIM, S. H., PARK, J., SON, Y., YOO, H. S., LEE, K. Y. & JANG, Y. S. 2013. Nasal immunization with major epitope-containing ApxIIA toxin fragment induces protective immunity against challenge infection with *Actinobacillus pleuropneumoniae* in a murine model. *Vet Immunol Immunopathol*, 151, 102-12.

SHINE, J. & DALGARNO, L. 1974. The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci USA*, 71, 1342-6.

SHIRATAKI, C., SHOJI, O., TERADA, M., OZAKI, S., SUGIMOTO, H., SHIRO, Y. & WATANABE, Y. 2014. Inhibition of heme uptake in *Pseudomonas aeruginosa* by its hemophore (HasA(p)) bound to synthetic metal complexes. *Angew Chem Int Ed Engl*, 53, 2862-6.

SMALLEY, J. W., BYRNE, D. P., BIRSS, A. J., WOJTOWICZ, H., SROKA, A., POTEMPA, J. & OLCZAK, T. 2011. HmuY haemophore and gingipain proteases constitute a unique syntrophic system of haem acquisition by *Porphyromonas gingivalis*. *PLoS One*, 6, e17182.

SMALLEY, J. W., CHARALABOUS, P., BIRSS, A. J. & HART, C. A. 2001. Detection of heme-binding proteins in epidemic strains of *Burkholderia cepacia*. *Clin Diagn Lab Immunol*, 8, 509-14.

SPOLAORE, B., DE FILIPPIS, V. & FONTANA, A. 2005. Heme binding by the N-terminal fragment 1-44 of human growth hormone. *Biochemistry*, 44, 16079-89.

SPRENCEL, C., CAO, Z., QI, Z., SCOTT, D. C., MONTAGUE, M. A., IVANOFF, N., XU, J., RAYMOND, K. M., NEWTON, S. M. & KLEBBA, P. E. 2000. Binding of ferric enterobactin by the *Escherichia coli* periplasmic protein FepB. *J Bacteriol*, 182, 5359-64.

SRIKUMAR, R., MIKAEL, L. G., PAWELEK, P. D., KHAMESSAN, A., GIBBS, B. F., JACQUES, M. & COULTON, J. W. 2004. Molecular cloning of haemoglobin-binding protein HgbA in the outer membrane of *Actinobacillus pleuropneumoniae*. *Microbiology*, 150, 1723-34.

SRINIVAS, V. & RAO, C. M. 1990. Time profile of hemin aggregation: an analysis. *Biochem Int*, 21, 849-55.

STELLARI, F. F., LAVRENTIADOU, S., RUSCITTI, F., JACCA, S., FRANCESCHI, V., CIVELLI, M., CARNINI, C., VILLETTI, G. & DONOFRIO, G. 2014. Enlightened *Mannhemia haemolytica* lung inflammation in bovinized mice. *Vet Res*, 45, 8.

STOJILJKOVIC, I. & HANTKE, K. 1995. Functional domains of the *Escherichia coli* ferric uptake regulator protein (Fur). *Mol Gen Genet*, 247, 199-205.

STOKES, R. H., OAKHILL, J. S., JOANNOU, C. L., GORRINGE, A. R. & EVANS, R. W. 2005. Meningococcal transferrin-binding proteins A and B show cooperation in their binding kinetics for human transferrin. *Infect Immun*. 73 (2), 944-52.

STRANGE, H. R., ZOLA, T. A. & CORNELISSEN, C. N. 2011. The fbpABC operon is required for Ton-independent utilization of xenosiderophores by *Neisseria gonorrhoeae* strain FA19. *Infect Immun*. 79 (1), 267-78.

STRAUS, D. C., UNBEHAGEN, P. J. & PURDY, C. W. 1993. Neuraminidase production by a *Pasteurella haemolytica* A1 strain associated with bovine pneumonia. *Infect Immun*, 61, 253-9.

SYLTE, M. J., CORBEIL, L. B., INZANA, T. J. & CZUPRYNSKI, C. J. 2001. *Haemophilus somnus* induces apoptosis in bovine endothelial cells in vitro. *Infect Immun*, 69, 1650-60.

TATUM, F. M., BRIGGS, R. E., SREEVATSAN, S. S., ZEHR, E. S., LING HSUAN, S., WHITELEY, L. O., AMES, T. R. & MAHESWARAN, S. K. 1998. Construction of an isogenic leukotoxin deletion mutant of *Pasteurella haemolytica* serotype 1: characterization and virulence. *Microb Pathog*, 24, 37-46.

TAUSEEF, I., ALI, Y. M. & BAYLISS, C. D. 2013. Phase variation of PorA, a major outer membrane protein, mediates escape of bactericidal antibodies by *Neisseria meningitidis*. *Infect Immun*, 81, 1374-80.

TAUSEEF, I., HARRISON, O. B., WOOLDRIDGE, K. G., FEAVERS, I. M., NEAL, K. R., GRAY, S. J., KRIZ, P., TURNER, D. P., ALA'ALDEEN, D. A., MAIDEN, M. C. & BAYLISS, C. D. 2011. Influence of the combination and phase variation status of the haemoglobin receptors HmbR and HpuAB on meningococcal virulence. *Microbiology*. 157 (5), 1446-56.

TAYLOR, J, D., FULTON, R. W., LEHENBAUER, T. W., STEP, D. L., & CONFER, A. W. 2010. The epidemiology of bovine respiratory disease: What is the evidence for preventive measures? *Can Vet J.* 51(12) 1351–1359.

THUMBIKAT, P., BRIGGS, R.E., KANNAN, M.S. & MAHESWARAN, S.K., 2003. Biological effects of two genetically defined leukotoxin mutants of *Mannheimia haemolytica*. *Microbial Pathogenesis,* 34 (5), 217-226

USDA. 2011. Researchers Uncover Genetic Link to Cattle Diseases [Online]. United States Department of Agriculture. Available: http://www.ars.usda.gov/is/pr/2011/110908.htm [Accessed 19/07/2012.

VAN DONKERSGOED, J. 1992. Meta-analysis of field trials of antimicrobial mass medication for prophylaxis of bovine respiratory disease in feedlot cattle. *Can Vet J*, 33, 786-95.

VAN OVERBEKE, I., CHIERS, K., CHARLIER, G., VANDENBERGHE, I., VAN BEEUMEN, J., DUCATELLE, R. & HAESEBROUCK, F. 2002. Characterization of the in vitro adhesion of *Actinobacillus pleuropneumoniae* to swine alveolar epithelial cells. *Vet Microbiol*, 88, 59-74.

VAN STOKKUM, I. H., SPOELDER, H. J., BLOEMENDAL, M., VAN GRONDELLE, R. & GROEN, F. C. 1990. Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Anal Biochem*. 191 (1), 110-8.

WEIRZBICKA, M. 2014. Studies on Structure, Function and Immunogenicity of Neisserial Secreted Protein Gly1ORF1. PhD Thesis, The University of Sheffield.

WHITBY, P. W., VANWAGONER, T. M., SEALE, T. W., MORTON, D. J. & STULL, T. L. 2006. Transcriptional profile of *Haemophilus influenzae*: effects of iron and heme. *J Bacteriol*. 188 (5), 5640-5.

WHITE, D. C. & GRANICK, S. 1963. Hemin biosynthesis in Haemophilus. *J Bacteriol*, 85, 842-50.

WHITMORE, L. & WALLACE, B. A. 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers*, 89, 392-400.

WHO 2012. Global incidence and prevalence of selected curable sexually transmitted infections - 2008. World Health Organisation, Dept. of Reproductive Health and Research.

WOLFF, N., IZADI-PRUNEYRE, N., COUPRIE, J., HABECK, M., LINGE, J., RIEPING, W., WANDERSMAN, C., NILGES, M., DELEPIERRE, M. & LECROISEY, A. 2008. Comparative analysis of structural and dynamic properties of the loaded and unloaded hemophore HasA: functional implications. *J Mol Biol*, 376, 517-25.

WOO, J. S., ZELTINA, A., GOETZ, B. A. & LOCHER, K. P. 2012. X-ray structure of the *Yersinia pestis* heme transporter HmuUV. *Nat Struct Mol Biol*. 19 (12), 1310-5.

WYCKOFF, E. E., DUNCAN, D., TORRES, A. G., MILLS, M., MAASE, K. & PAYNE, S. M. 1998. Structure of the *Shigella dysenteriae* haem transport locus and its phylogenetic distribution in enteric bacteria. *Mol Microbiol*, 28, 1139-52.

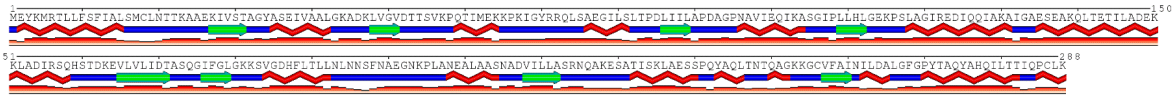
YANG, X., YU, R. H., CALMETTES, C., MORAES, T. F. & SCHRYVERS, A. B. 2011. The Anchor Peptide of Transferrin Binding Protein B is Required for Interaction with Transferrin Binding Protein A. *J Biol Chem*. 286 (52), 45165–45173.

YEE, V. C., PEDERSEN, L. C., LE TRONG, I., BISHOP, P. D., STENKAMP, R. E. & TELLER, D. C. 1994. Three-dimensional structure of a transglutaminase: human blood coagulation factor XIII. *Proc Natl Acad Sci U S A*, 91, 7296-300.

YU, C. & GENCO, C. A. 2012. Fur-mediated activation of gene transcription in the human pathogen *Neisseria gonorrhoeae*. *J Bacteriol*. 194 (7), 1730-42.

YU, N. Y., WAGNER, J. R., LAIRD, M. R., MELLI, G., REY, S., LO, R., DAO, P., SAHINALP, S. C., ESTER, M., FOSTER, L. J. & BRINKMAN, F. S. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*. 26 (13), 1608-15.

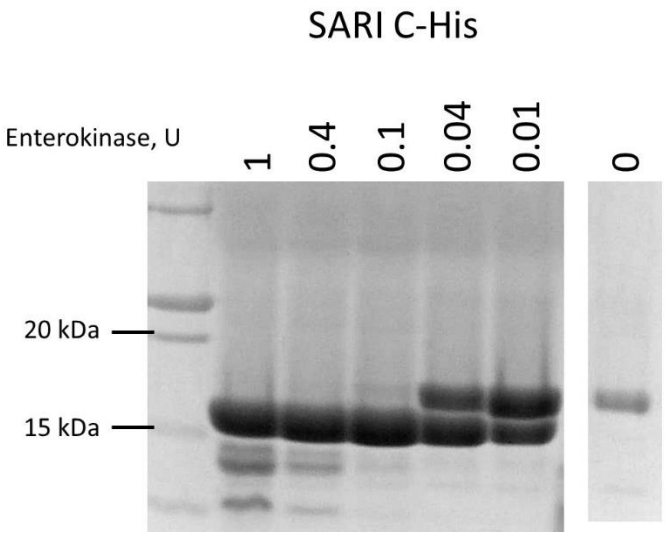
9. Appendix



Appendix 1. PhuT secondary structure as predicted by SABLE server.

Cell type	Genotype
K-12 M72	lac(am) trp(am) rpsL lambda cI857 deltaH1 bio252
BL21(DE3)	F– ompT hsdSB(rB–, mB–) gal dcm (DE3)
BL21(DE3)RIPL	B F– ompT hsdS(rB– mB–) dcm+Tetrgalλ(DE3) endA Hte [argU proLCamr] [argU ileY leuW Strep/Specr]
XLI Blue	endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(rK- mK+)

Appendix 2. Plasmids used within this thesis.



Appendix 3. Enterokinase cleavage of SARI C-His polyhistidine tag.