UNIVERSITY
of
GLASGOW

Comparative outer membrane proteomics of *Pasteurella multocida* isolates associated with diseased cattle, sheep, pigs and chickens

Teerasak E-komon

Presented for the degree of Doctor of Philosophy in the Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow

October 2011

Declaration

I hereby declare that the work presented in this thesis is my own, except where otherwise cited or acknowledged. No part of this thesis has been presented for any other degree. The research for this thesis was performed between November 2007 and September 2011.

Teerasak E-komon

Acknowledgement

I would most like to thank my supervisors: Dr. Robert Davies, Dr. Richard Burchmore and Dr. Pawel Herzyk for their continuous supervision, advice, support and encouragement throughout the course of my study and also for their assistance and patience on reading and commenting on this thesis.

Special thanks to Prof. John Coote and Dr. Daniel Walker as my assessors for their encouragement and helpful suggestions.

I would like to thank everybody in the Proteomics group, Glasgow Polyomics Facility, School of Life Sciences, University of Glasgow for their guidance, materials contribution, collaboration and helpful support during this study.

Many thanks also go to Mrs. Susan Bailie, Mrs. June Southall, Mrs. Denise Candlish, Mrs. Mulu Gedle and all members of Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow for their technical help, collaboration and friendship.

Finally, I would like to thank the Thai Ministry of Higher Education and the Office of Civil Service Commission for granting this study.

Abstract

The Gram-negative bacterial pathogen *Pasteurella multocida* causes economically significant infections of domesticated animals. Very little is known about the roles of *P. multocida* outer membrane proteins (OMPs) in host-specificity and virulence. This study aimed to compare the outer membrane proteomes of eight representative *P. multocida* isolates associated with diseased cattle (two), sheep (two), pigs (two) and poultry (two).

Ten different predictors classified into three groups (subcellular localization, transmembrane β-barrel protein and lipoprotein predictors) were used to identify putative OMPs from two available *P. multocida* genomes: those of avian strain Pm70 and porcine non-toxigenic strain 3480. Predicted proteins in each group were filtered by optimized criteria for consensus prediction: at least two positive predictions for the subcellular localization predictors, three for the transmembrane β-barrel protein predictors and one for the lipoprotein predictors. The consensus predicted proteins were integrated from each group into a single list of proteins. This study further incorporated a manual confirmation step including a public database search against PubMed and sequence analyses, e.g. sequence and structural homology, conserved motifs/domains, functional prediction, and protein-protein interactions to enhance the confidence of prediction. Filtered-out proteins were analysed by manual confirmation. As a result, we were able to confidently predict 105 putative OMPs from the avian strain genome and 107 OMPs from the porcine strain genome with 83% overlap between the two genomes.

By using a combination of gel-based and gel-free proteomic methods, outer membrane peptides were obtained by in-gel and in-solution tryptic digests of Sarkosyl-extracted OMPs and identified by MALDI-TOF-TOF MS and LC-ESI-Q/TOF MS. Fifty-four

different OMPs were detected and these represented 52% of the predicted avian outer membrane sub-proteome and 48% of the predicted porcine sub-proteome. Twenty-four core proteins, involved mainly in outer membrane biogenesis and integrity, or having transport and receptor functions, were identified in isolates from all four animal hosts. Conversely, other proteins with functions primarily in adherence and colonization, or as TonB-dependent iron receptors, were restricted to only one or a few isolates. Proteomic analysis of the cell envelope profiles of the same isolates identified 10 proteins that had been lost during Sarkosyl extraction. Thus, in total, 64 OMPs were identified among the eight isolates and these represented 62% of the predicted avian outer membrane sub-proteome and 57% of the predicted porcine sub-proteome. Thirty-six of these were core OMPs and 28 proteins were restricted to certain isolates or to some animal hosts.

Outer membrane proteomes of the eight isolates were compared after growth under different growth conditions using a combination of gel-based and gel-free methods. Bacteria were harvested at different stages of the growth, grown under different rates of aeration, under iron-replete and iron-restricted conditions, in different sera, in various culture media supplemented with different sera, and on solid surfaces as biofilms. Slight changes were observed in the OMP profiles at different stages of the growth. Different rates of aeration affected the expression of iron receptor proteins. High aeration reduced the expression of iron receptors, whereas low aeration increased the expression of these proteins. Iron receptor proteins were highly expressed in all of the isolates grown under iron-limited conditions. HgbA and HemR were expressed in all of the isolates grown under iron-limited conditions whereas, HasR, TbpA, PfhR, two HgbB proteins, Hup, and TonB-dependent receptors PM0803, PM1428, PM0741 and PM1282 were expressed in some isolates. The effect of growth in serum on the OMP profiles was dependent on serum composition. The OMP profiles obtained after growth in the tissue culture medium M199 resembled those grown under iron-limited conditions. The effect of adding serum to

M199 on the OMP profiles depended on sera. Adding chicken and foetal calf sera reduced expression of iron receptor proteins compared to growth in M199 alone. Growth on different types of agar media including brain heart infusion (BHI) agar, BHI agar supplemented with sheep's blood, BHI agar supplemented with sucrose and BHI agar supplemented with sucrose and Congo Red resulted in changes to the OMP profiles in comparison to growth in broth. Core OMPs (e.g., OmpA, OmpH, OmpP6/Pal, FadL and Oma87) were expressed in all isolates and under all growth conditions, whereas the remainder (e.g., TbpA, TadD, RcpA, LppB/NlpD, OmpLA, HgbA, HbpA, HasR, HmbR, SrfB, Wza, LspB) were expressed in certain isolates under different growth conditions. The expression of some OMPs (Opa, Hsf and NanH) was restricted to the same isolates under different growth conditions. Opa was expressed under all growth conditions in only avian isolates; Hsf was expressed in avian isolates and ovine isolate PM966 under growth in M199 and M199 supplemented with serum and on agar media; NanH was expressed in bovine isolate PM632 under all growth conditions.

This study represents the first comparative outer membrane proteomic analysis after growth of *P. multocida* isolates associated with diseases from different animal hosts under different growth conditions.

Table of contents

List of tables

List of figures

List of accompanying material

Appendix Table 2-1. Proteins predicted from the avian strain genome of *P. multocida* using 10 bioinformatic predictors

Appendix Table 2-2. Proteins predicted from the porcine strain genome of *P. multocida* using 10 bioinformatic predictors

Appendix Table 3-1. Proteins identified from the OMP samples extracted by using different concentration of Sarkosyl (0.1%, 0.5%, 1% and 2%).

Appendix Table 3-2. Proteins identified from the OMP samples extracted by using different detergents.

Appendix Table 3-3. OMPs identified by gel-based proteomics

Appendix Table 3-4. Additional OMPs identified by gel-based proteomics followed by LC-MS/MS

Appendix Table 3-5. OMPs identified by gel-free proteomics

Appendix Table 3-6. Identified proteins obtained from extraction of cell envelopes by gel-based and gel-free proteomics

Appendix Table 4-1. Proteins identified from growth of eight isolates of *P. multocida* in sera by gel-based and gel-free proteomics

Appendix Table 4-2. Proteins identified from growth of eight isolates of *P. multocida* in M199 and in M199 supplemented with sera by gel-based and gel-free proteomics

Appendix Table 4-3. Proteins identified from growth of eight isolates of *P. multocida* on different agar media by gel-based and gel-free proteomics

Abbreviations

1D	=	One dimensional
2D	=	Two dimensional
BHI	=	Brain heart infusion
CAI	=	Codon adaptation index
DNA	=	Deoxyribonucleic acid
DNT	=	Dermonecrotic toxin
DTT	=	Dithiothreitol
ESI	=	Electrospray ionization
GRAVY	=	Grand average of hydrophobicity
h	=	Hour
HMM	=	Hidden Markov Model
HMMTOP	=	Hidden Markov Model for Topology Prediction
HS	=	Haemorrhagic septicaemia
IAA	=	Iodoacetamide
IM	=	Inner membrane
kDa	=	Kilodalton
LC	=	Liquid chromatography
LOS	=	Lipooligosaccharide
LPS	=	Lipopolysaccharide
MALDI	=	Matrix-assisted laser desorption/ionization
min	=	Minute
MLST	=	Multilocus sequence typing
mRNA	=	messenger RNA
MS	=	Mass spectrometry
MW	=	Molecular weight
NN	=	Neural Network
ºC	=	Degree celsius
OM	=	Outer membrane
OMP	=	Outer membrane protein
ORF	=	Open reading frame
PAR	=	Progressive atrophic rhinitis
Pcho	=	Phosphocholine
PCR	=	Polymerase chain reaction
p*I*	=	Isoelectric pH
PM	=	*Pasteurella multocida*
PMT	=	*P. multocida* toxin
rRNA	=	ribosomal RNA
SDS-PAGE	=	Sodium dodexyl sulphate - polyacrylamide gel electrophoresis
SVM	=	Support Vector Machine
TOF	=	Time-of-flight
V	=	Volt
v/v	=	Volume/volume
w/v	=	Weight/volume

Chapter 1: General introduction

1.1 Classification

1.1.1 The family Pasteurellaceae

The family *Pasteurellaceae* is a large and diverse group of chemoorganotrophic, facultatively anaerobic and fermentative Gram-negative Proteobacteria and comprises approximately sixteen genera: *Pasteurella*, *Actinobacillus*, *Haemophilus, Aggregatibacter, Avibacterium, Bibersteinia, Bisgaardia, Gallibacterium, Histophilus, Lonepinella, Mannheimia, Nicoletella, Phocoenobacter, Terrahaemophilus, Volucribacter* and unclassified *Pasteurellaceae* (Wheeler *et al.*, 2005). They live as commensals in vertebrate (mammals, birds and reptiles) species and many of them are important veterinary pathogens.

1.1.2 The genus Pasteurella

The genus *Pasteurella*, which was named to commemorate Louis Pasteur's work, was first described in association with an outbreak of fowl cholera (Christensen & Bisgaard, 2006). *Pasteurella multocida* was the first designated and pivotal species of this genus. Other members were described based on phenotypic similarities such as *Pasteurella gallinarum*, *Pasteurella dagmatis*, *Pasteurella canis*, *Pasteurella stomatis*, *Pasteurella avium*, *Pasteurella volantium*, *Pasteurella langaa*, *Pasteurella anatis* and two unnamed taxa (*Pasteurella* species A and B). The genus can be differentiated from other genera on the basis of a lack β-haemolysis activity and an ability to ferment D-fructose, D-galactose, D-mannose and sucrose, but not glycosides and D-melibiose (Christensen & Bisgaard, 2006).

1.1.3 Pasteurella multocida

P. multocida is a commensal and pathogenic bacterium in many mammals and birds. *P. multocida* is different from other members of the *Pasteurella* genus because it gives

positive reactions for ornithine decarboxylase and can ferment D-mannitol, but not maltose and dextrin (Christensen & Bisgaard, 2006). The closest taxon of *P. multocida* is *P. dagmatis* (Christensen & Bisgaard, 2006). On the basis of DNA comparison using DNA-DNA hybridization, *P. multocida* is classified into four subspecies **(Figure 1-1)**: *P. multocida* subsp. *multocida*, *P. multocida* subsp. *septica*, *P. multocida* subsp. *gallicida* (Mutters *et al.*, 1985; Adlam & Rutter, 1989) and the recently described *P. multocida* subsp. *tigris* (Harper *et al.*, 2006). The first three subspecies can be separated by differences in fermentation of D-sorbitol and dulcitol (Adlam & Rutter, 1989; Donachie *et al.*, 1995).

1.1.4 Cell shape and colony morphology

P. multocida is a nonmotile coccobacillus bacterium which grows well on enriched agar media supplemented with 5% inactivated serum or blood from cattle, horse or sheep at 35-37^{o}C for 18-24 h. However, an optimal growth temperature for the avian strains might be as high as 42ºC. Under these conditions, colonies may range in size from 1 to 3 mm in diameter. Variation in colony morphology can be observed in different strains ranging from mucoid to smooth forms with a sweetish smell of indole (Christensen & Bisgaard, 2006). Mucoid colonies are composed of cells with capsules consisting in part of hyaluronic acid. Large watery mucoid colonies **(Figure 1-2A)** are associated with isolates from ruminants, pigs and rabbits (Christensen & Bisgaard, 2006). Smooth colonies **(Figure 1-2B)** are the combination of capsulated and noncapsulated strains. Conversely, rough colonies comprise filamentous noncapsulated strains and are not commonly found. Generally, the colonies of noncapsulated isolates are not iridescent and appear blue, greyish-blue or grey. On the other hand, capsulated strains have yellowish-green, bluish-green or pearl-like iridescence. There is a relationship between capsular type and colony morphology. All serotype A and a few serotype D strains have mucoid colonies. Pearl-

Superkingdom Bacteria
Phylum Proteobacteria
Class Gammaproteobacteria
Order Pasteurellales
Family *Pasteurellaceae*
Genus *Pasteurella*
Species *multocida*
Subspecies g*allicida, multocida, septica* and *tigris*

Figure 1-1. Taxonomy of *P. multocida* (Wheeler *et al.*, 2005)

A) **B)**

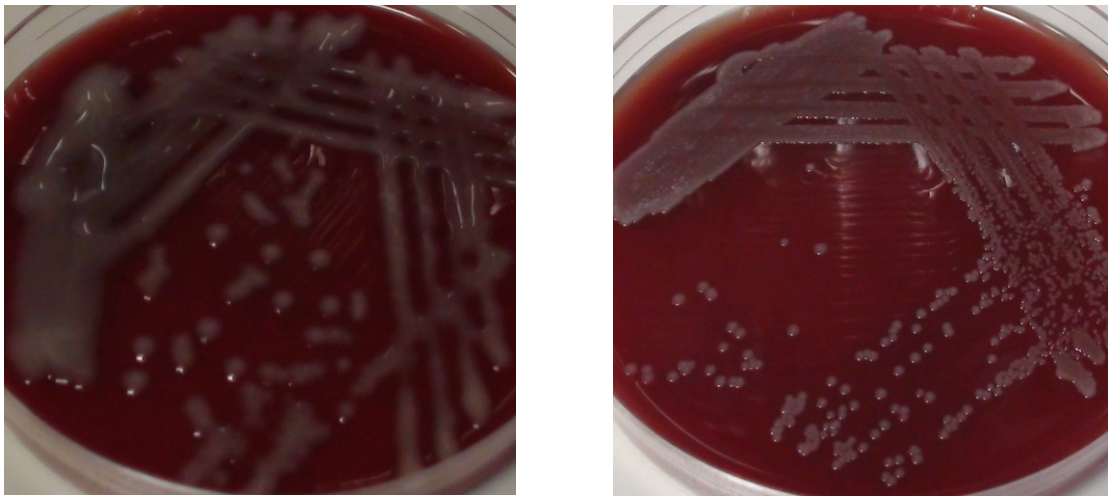


Figure 1-2. Comparison of watery mucoid **(A)** and smooth or non-mucoid **(B)** colony morphologies of *P. multocida* strains grown on Brain Heart Infusion agar supplemented with 5% defibrinated sheep blood.

like iridescent colonies belong to serotypes A, D and F. Large watery mucoid colonies are produced by serotype A and smooth yellowish- or bluish-green iridescent colonies usually belong to serotype B and E (Adlam & Rutter, 1989).

1.2 Typing methods

To better understand the epidemiology and host predilection of *P. multocida*, systematic classification into subspecies or strains is important. Four main approaches that have been used are described below.

1.2.1 Biotyping

Biochemical tests are the traditional method of classifying *P. multocida* into biotypes or subspecies because of variation in the utilization of sugars such as pentoses, disaccharides and polyhydric alcohols, but it is often not possible to clearly identify different strains (Adlam & Rutter, 1989). *P. multocida* subsp. *multocida* is dulcitol-negative and sorbitol-positive, whereas *P. multocida* subsp. *septica* is dulcitol-negative and sorbitol-negative. *P. multocida* subsp. *gallicida* is dulcitol-positive and sorbital-positive (Petersen *et al.*, 2001). Similar to *P. multocida* subsp. *multocida*, *P. multocida* subsp. *tigris* is dulcitol-negative and sorbitol-positive, but is differentiated from subsp. *multocida* by DNA analysis (Capitini *et al.*, 2002).

1.2.2 Serological typing

Several techniques have been used in serological studies of *P. multocida* such as agglutination and adsorption tests, passive haemagglutination, passive protection of mice and agar gel diffusion precipitation (Dziva *et al.*, 2008). The Carter system classified *P. multocida* into 5 capsular serogroups (A, B, D, E and F) based on passive haemagglutination of erythrocytes sentisized by specific capsule antigen (Adlam & Rutter, 1989). Some strains of *P. multocida* are untypable. Generally, serotype B and E strains are associated with haemorrhagic septicaemia (HS) cases in ruminants and pigs whilst

serotypes A and F are commonly recovered from cases of avian fowl cholera (Townsend *et al.*, 1998). Serotype A strains are associated with pneumonia in ruminants and pigs. Cases of atrophic rhinitis are associated with serotype D and A. However, the use of these associations as an indicator of host predilection is unsafe because certain observations have revealed possible changing epidemiology of *P. multocida* (Dziva *et al.*, 2008). Alternatively, somatic lipopolysaccharide (LPS) typing can be used for the identification of *P. multocida* (Adlam & Rutter, 1989). There are two main systems reported and these were mainly based on the avian isolates. The Namioka system is based on a tube agglutination test and is able to distinguish 11 serotypes, while the Heddleston system is based on a gel-diffusion precipitation test and can recognize 16 serotypes (Adlam & Rutter, 1989). The latter system is currently the preferred method. A standard system for the identification of *P. multocida* serotypes has been recommended to utilise both the Carter capsular system identified by letters (A, B, D, E and F) and the Heddleston somatic typing system identified by numbers (1-16) (Adlam & Rutter, 1989; St Michael *et al.*, 2005).

1.2.3 Macromolecular profiling

Alternatively, outer membrane protein (OMP) profiling based on electrophoretic migration of the OMPs confers patterns for strain typing (Dziva *et al.*, 2008). Lugtenberg *et al.* (1984) established a correlation between OMP profiles and the pathogenicity of strains causing swine atrophic rhinitis predicted by the guinea pig skin test. Davies *et al.* (2003a, 2003b, 2004) used the OMP typing to classify *P. multocida* strains isolated from different animal species. OMP typing scheme was devised-based, firstly, on molecular variation of the two major proteins, OmpA and OmpH (OMP-type 1, 2, etc.), and, secondly, on variation of minor protein patterns (OMP-type 1.1, 2.1, etc). The electrophoretic patterns of the major OMPs, OmpH and a heat modifiable OmpA, and other minor proteins divided avian strains of *P. multocida* into 19 OMP types and revealed relatively high diversity of

the avian strains (Davies *et al.*, 2003a). Conversely, bovine strains showed lower diversity (Davies *et al.*, 2004). Similarly, OMP profiling in association with capsule typing and the presence or absence of the *toxA* gene showed that different porcine strains of *P. multocida* were responsible for pneumonia and atrophic rhinitis (Davies *et al.*, 2003b). Different patterns of OMP profiles from *P. multocida* strains may indicate different modes of host-pathogen interaction. However, correlation of OMP types with disease-status, host species and geographic area is not fully determined and there is a possibility that different proteins such as TolC and FadL with similar electrophoretic mobility provide the same OMP pattern. Contamination of proteins from other subcellular locations can interfere with the interpretation of the OMP profiles.

1.2.4 Genotyping methods

Various DNA-based techniques are used in the molecular typing of different isolates of *P. multocida*. Restriction endonuclease analysis (REA) generates specific patterns based on DNA cleavage with a restriction enzyme such as *Bgl*II, *Hpa*II or *Hha*I (Christensen & Bisgaard, 2006). In ribotyping, restriction enzyme-digestion products are transferred onto a nitrocellulose membrane which is hybridized with 16S or 23S rRNA probes (Christensen & Bisgaard, 2006).

PCR-based methods are applied as a tool for the rapid and specific detection of microorganisms and the diagnosis of disease by using specific primers tracking a conserved and unique gene within the bacterial genome (Dziva *et al.*, 2008). As *P. multocida* is the causative agent of different diseases in various hosts, the application of PCR technology for disease-specific strains is very significant (Hunt *et al.*, 2000; Dziva *et al.*, 2008). Miflin and Blackall (2001) accurately tested unrelated avian and porcine strains of *P. multocida* with primers targeting the 23S rRNA gene. Lichtensteiger *et al.* (1996) developed PCR assays targeting a fragment of the *toxA* gene, which encodes a *P.*

multocida toxin for atrophic rhinitis in swine, to discriminate between toxigenic and nontoxigenic strains of *P. multocida*. Choi and Chae (2001) used nested PCR to identify the *toxA* gene and differentiate between porcine toxigenic and non-toxigenic strain of *P. multocida*. Brickell *et al.* (1998) developed PCR assays using 16S rRNA-23S rRNA products to uniquely identify HS-causing strains of *P. multocida* serotype B:2 in cattle and buffalo. PCR primers specific to the hyaluronic acid encoding region (hyaC-hyaD) of *P. multocida* provided successful identification of fowl cholera caused by avian *P. multocida* strains of serogroup A (Gautam *et al.*, 2004). Davies *et al.* (2003a; 2003b; 2003c; 2004) determined the capsular types of avian, bovine, porcine and ovine *P. multocida* isolates by multiplex capsular PCR typing. Multiplex PCR, capsular typing was used to identify capsular types A, D and F from Indian avian strains (Shivachandra *et al.*, 2006).

Another aspect of the genotype-based approach is DNA sequence comparison which can be used in the classification of *P. multocida* at species and subspecies levels (Dziva *et al.*, 2008). Davies *et al.* (2004) inferred that bovine *P. multocida* strains had a low degree of genetic diversity using multilocus sequence analysis (MLSA) with DNA sequences from seven housekeeping genes (*adk*, *aroA*, *deoD*, *gdhA*, *g6pd*, *mdh* and *pgi*). Similarly, Kuhnert & Korczak (2006) used sequencing of three genes (*recN*, *rpoA* and *thdF*) to study the phylogeny of the family *Pasteurellaceae*.

DNA-DNA hybridisation is another application for the diagnosis of *P. multocida*. Register *et al.* (1998) used probes derived from the *P. multocida toxA* gene to detect toxigenic *P. multocida* strains causing atrophic rhinitis in swine by colony hybridization. Likewise, fluorescent *in situ* hybridization (FISH) with probes designed to target 16S rRNA detected *P. multocida* strains causing fowl cholera in chicken and respiratory tract infections in swine (Mbuthia *et al.*, 2001). These studies show that PCR products conjugated to specific compounds can easily be utilized as probes for hybridization assays.

Other techniques that have been used for typing are random amplification of polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) (Shivachandra *et al.*, 2006b), repetitive extragenetic palindromic (REP)-PCR and enterobacterial repetitive intergenic consensus (ERIC)-PCR (Hunt *et al.*, 2000; Dziva *et al.*, 2008; Shivachandra *et al.*, 2008).

Recent multilocus sequence typing (MLST) study (**Figure 1-3**) of seven house-keeping enzyme genes from 119 isolates of *P. multocida* of bovine, ovine, porcine and avian origin showed clustering of different isolates of *P. multocida* in association with different diseases (Davies *et al.* unpublished data; http://pubmlst.org/pmultocia_multihost/). The majority of the bovine isolates and certain porcine isolates were grouped in the major bovine pneumonia cluster. The porcine isolates and certain bovine and avian isolates were grouped in the major porcine pneumonia cluster. The porcine isolates of capsular type D and certain avian isolates were grouped in the serotype D porcine atrophic rhinitis cluster. The majority of the ovine isolates were clustered in the avian/ovine serogroup F cluster. The avian/bovine/porcine/ovine isolate cluster contained isolates from all four animal hosts. This MLST study showed evidence of the association between different serotypes and diseases. The clusters of isolates associated with different animal hosts could be due to transmission of *P. multocida* from one host to another.

1.3 Diversity of *P. multocida* populations

P. multocida is a very diverse species which lives as a commensal or causes different diseases in many mammals and birds. This section reviews the diversity of *P. multocida* populations in different animal hosts.

1.3.1 Avian strains

Many studies have shown that avian strains of *P. multocida* are extremely diverse (Rhoades & Rimler, 1987; Hirsh *et al.*, 1990; Blackall *et al.*, 1998; Gunawardana *et al.*, 2000; Davies *et al.*, 2003a; Shivachandra *et al.*, 2006). Davies *et al.* (2003a) characterized 100 avian *P. multocida* isolates recovered from different disease cases (e.g. fowl cholera, septicaemia, sinusitis, conjunctivitis, oedema, swollen head and pericarditis) in England and Wales by using capsular PCR typing and comparison of OMP profiles. The authors suggested that the population structure of these isolates was clonal. The majority (68 percent) of these isolates were of capsular type A, whereas the remainder were of types B, D and F. These 100 isolates were also classified into 19 OMP-types. Fifteen OMP-types accounted for 56 percent of the isolates, whereas the other four types accounted for 44 percent. The high degree of diversity associated with a large number of disease cases suggested that they were opportunistic pathogens and possess a range of virulence factors. Moreover, the association of isolates of capsular types B, D and F with specific OMP-types suggested the evidence of distinct clonal populations. Similar observations were described by Rhoades and Rimler (1987) who determined the capsular and somatic serotypes of 246 avian isolates. These authors showed that 166 isolates were of capsular type A and 12 somatic serotypes. Jabbari *et al.* (2006) showed that all 39 isolates recovered from avian sources in Iran were of capsular type A.

1.3.2 Bovine strains

Bovine isolates are associated with two different types of disease; pneumonia which is caused worldwide by isolates of capsular type A and haemorrhagic septicaemia which is caused by isolates of capsular types B and E in Asia and Africa (Dabo *et al.*, 2007; Kumar *et al.*, 2009). For strains causing bovine pneumonia, Davies *et al.* (2004) used capsular PCR typing, analysis of OMP profiles and multilocus sequence comparison to study 153

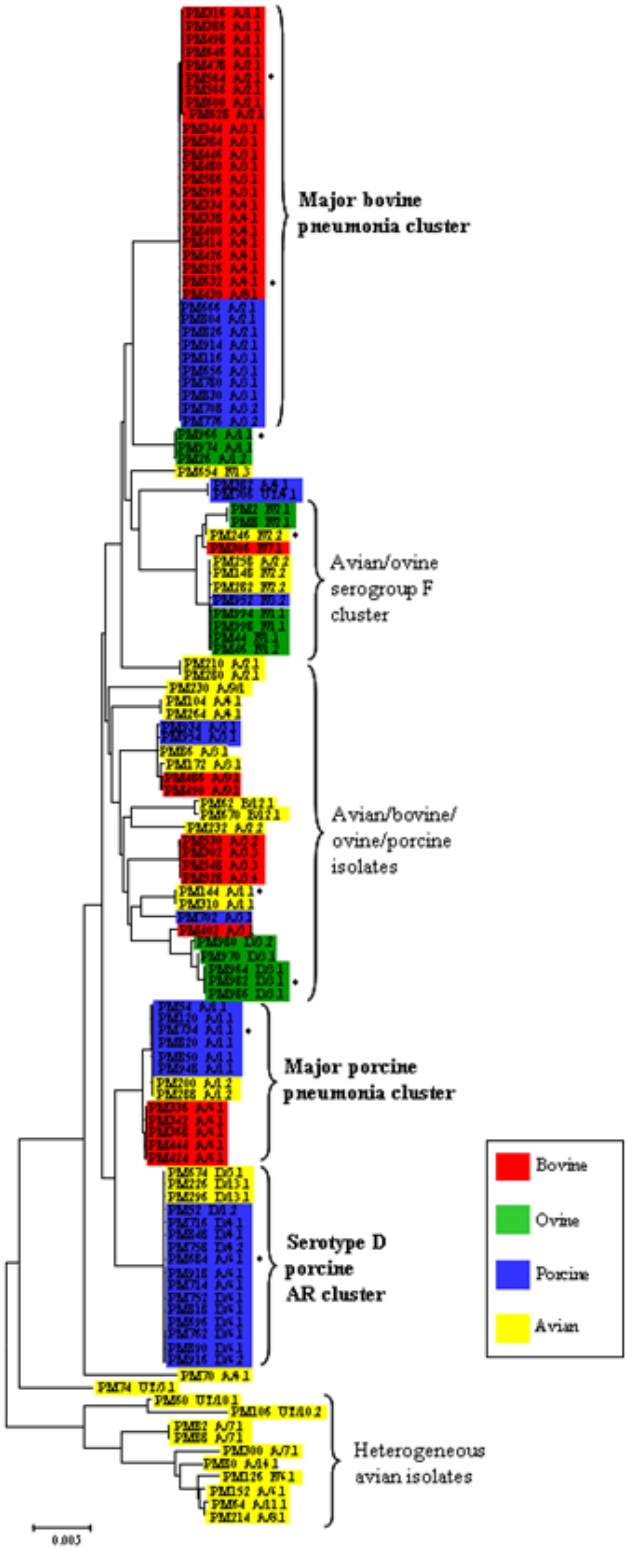


Figure 1-3. Neighbour-joining tree constructed from the concatenated sequences (3990 bp) of seven house-keeping enzyme genes from 119 isolates of *P. multocida* of bovine, ovine, porcine and avian origin (Davies *et al.* unpublished data). * indicates isolates used in the present study.

bovine isolates of *P. multocida* recovered from England and Wales. The authors found that these bovine strains had limited diversity and 99% of these strains were capsular type A. Based on the OMP-types, five out of 13 OMP-types represented 85% of the isolates, suggesting an increased ability of these clones to cause disease. Ewers *et al.* (2006) similarly reported that 92% of bovine isolates recovered from healthy and pneumonia cases belonged to capsular type A. Strains recovered from haemorrhagic septicaemia cases also showed limited diversity (Townsend *et al.*, 1997; Karunakaran *et al.*, 2009). Townsend *et al.* (1997) characterized 38 haemorrhagic septicaemia-causing isolates of *P. multocida* using repetitive extragenic palindromic (REP) PCR and demonstrated a high level of homogeneity among the isolates of capsular types B and E.

1.3.3 Porcine strains

Porcine isolates are associated with cases of progressive atrophic rhinitis (PAR) and pneumonia (Davies *et al.*, 2003b). Normally, PAR cases are caused by toxigenic isolates of capsular types D and A, whereas pneumonia cases are caused by non-toxigenic isolates of capsular type A (Choi & Chae, 2001; Ewers *et al.*, 2006; Ross, 2007; Tang *et al.*, 2009). Davies *et al.* (2003b) characterized 158 porcine isolates of *P. multocida* from cases of PAR and pneumonia in England and Wales using capsular PCR typing, expression of *toxA* gene and OMP profiles. The authors observed that the majority (76 percent) of cases of PAR were associated with toxigenic capsular type D (OMP-type 4.1) and toxigenic capsular type A and D (OMP-type 6.1) isolates. Similarly, the majority (88 percent) of cases of pneumonia were associated with non-toxigenic capsular type A (OMP-types 1.1, 2.1, 3.1 and 5.1) and non-toxigenic capsular type D (OMP-type 6.1) isolates. The limited diversity of these two subpopulations associated with PAR and pneumonia suggested that they were primary pathogens with a high level of virulence. These authors also demonstrated evidence of horizontal gene transfer between isolates, e.g. the presence or absence of the *toxA* gene within isolates of the same OMP-type 6.1. Another study by Djordjevic *et al.*

(1998) examined 22 Australian porcine isolates of *P. multocida* by using restriction endonuclease analysis. The authors showed that the toxigenic isolates of capsular type D associated with PAR cases had limited diversity, whereas the non-toxigenic isolates of capsular type A associated with pneumonia were more diverse.

1.3.4 Ovine strains

Ovine isolates of *P. multocida* are mostly associated with pneumonia (Davies *et al.*, 2003c; Odugbo *et al.*, 2006). Davies *et al.* (2003c) characterized 35 ovine isolates recovered from cases of pneumonia and neonatal septicaemia, and the vaginas of healthy ewes by using capsular PCR typing and analysis of OMP profiles. These authors classified the isolates into three capsular types (A, D and F) and three OMP-types, and identified four subpopulations. The cases of pneumonia were caused by different subpopulations compared to the cases of septicaemia and the isolates from the vaginal tracts of healthy ewes, suggesting adaptation of these isolates to different niches. Another study showed the presence of *toxA*-producing isolates of capsular types A and D in four healthy sheep (Shayegh *et al.*, 2010b). Similarly, Shayegh *et al.* (2008) also found that the majority of ovine isolates (39/47) were *toxA*-producing capsular type A. Weiser *et al.* (2003) analyzed 90 isolates associated with pneumonia in bighorn sheep using capsular and *toxA* PCR typing, and restriction fragment length polymorphism (RFLP) analysis. These authors observed diversity and showed a prevalence of non-toxigenic isolates of capsular type A. Kumar *et al.* (2009) also showed a prevalence of ovine isolates of capsular type A (26/28) associated with pneumonia in India.

1.3.5 Strains from other animal hosts

The population diversity of *P. multocida* strains has also been studied in several other animal hosts. Isolates of capsular type A were mostly found in cases of human respiratory tract infections resulting from scratches by cats and dogs (Donnio *et al.*, 2004). This was

consistent with the finding by Mohan *et al.* (1997) that isolates recovered from diseased cats and dogs (e.g. rhinitis, bronchitis, pneumonia, pyothorax, skin bite wounds and urogenital infections) in Zimbabwe belonged to capsular type A. A study in rabbits by Jaglic *et al.*(2005; 2006). showed that 17 of 27 rabbit isolates associated with rhinitis and pneumonia in the Czech Republic were of capsular type A, whereas the rest were of capsular types D (2) and F(8). Isolates of capsular type A have also been observed in goats, rabbits, leopards and deer in India (Kumar *et al.*, 2004). Saxena *et al.* (2006) have identified capsular types A and B in five Indian isolates obtained from lions and tigers.

1.4 Disease manifestations

P. multocida exists as a commensal in the upper respiratory tract of many animal hosts. Occasionally, isolation of *P. multocida* from the vagina of sheep, horses, dogs and rabbits has been reported (Watson & Davies, 2002). It can also cause diseases ranging from acute to chronic in a wide array of animal species including fowl cholera in poultry, atrophic rhinitis in swine, haemorrhagic septicaemia in cattle and water buffaloes, pneumonia in cattle, sheep and pigs, snuffles in rabbits, retropharyngeal infections in horses and wound abscesses and meningitis in humans resulting from cat and dog bites (Adlam & Rutter, 1989; Amory *et al.*, 2006). This section describes these diseases and **Table 1-1** shows a summary of diseases in different animal hosts caused by *P. multocida*.

1.4.1 Fowl cholera

Fowl cholera is an economically significant avian disease throughout the world (Adlam & Rutter, 1989). Capsular types A:1, A:3 and A:4 are a major cause of widely distributed fowl cholera in poultry (Adler *et al.*, 1999). A small number of infections are caused by capsular types B, D and F. Capsulated strains are more virulent than non-capsulated strains (Adlam & Rutter, 1989). All types of birds are susceptible to this disease but the degree of susceptibility varies (Adlam & Rutter, 1989). Transmission of this disease is

through exposure to water contaminated by infected birds (Adlam & Rutter, 1989). Common symptoms of acute fowl cholera are depression, ruffled feathers, fever, anorexia, oral mucous discharge, increased respiratory rate and diarrhoea. Infected birds show these symptoms in a few hours before death. Localized infections are usually associated with chronic infection, e.g. swelling wattles, sinuses, periorbital subcutaneous tissues, leg or wing joints, sterna bursea and foot pads, exudative conjunctivitis, pharyngitis, emaciation and lethargy (Adlam & Rutters, 1989).

1.4.2 Pneumonic pasteurellosis

1.4.2.1 Bovine pneumonic pasteurellosis

Pneumonic pasteurellosis is an economically important disease caused by virus, mycoplasma and bacterial infections including *Mannheimia haemolytica* and *P. multocida* (Adlam & Rutter, 1989). *P. multocida* capsular type A:3 strains cause pneumonic pasteurellosis or enzootic calf pneumonia in cattle (Adlam & Rutter, 1989). The cause of this disease is due to physical and emotional stress, and overcrowding (Adlam & Rutter, 1989). *P. multocida* is found in the nasal passages of both diseased and healthy cattle but not all cattle exposed to the bacteria will develop pneumonia (Dabo *et al.*, 2007). The common symptoms of infected cattle typically include depression, inappetence, cough, nasal discharge and fever (Dabo *et al.*, 2007).

1.4.2.2 Porcine pneumonic pasteurellosis

P. multocida causes porcine pneumonia as a secondary infection and isolates of capsular type A are usually responsible. The respiratory tract may become filled with frothy fluid, causing affected pigs to have difficulty breathing and to exhibit a characteristic “thumping” respiratory movement. Dry cough and increased body temperature are also symptoms (Adlam & Rutter, 1989). The disease can be acute or chronic lasting up to five weeks with relapses, weight loss and weakness (Adlam & Rutter, 1989).

1.4.2.3 Ovine pneumonic pasteurellosis

P. multocida also causes pneumonic pasteurellosis in sheep and an unusual septicaemia in neonatal lambs. Davies *et al.* (2003c) identified *P. multocida* strains of three capsular types (A, D and F) recovered from the vagina and respiratory tract of sheep. The symptoms of ovine pneumonic pasteurellosis include anorexia, coughing, nasal discharge without fever and dullness (Odugbo *et al.*, 2006).

1.4.3 Atrophic rhinitis

Capsular type A and D strains of *P. multocida* cause atrophic rhinitis; these strains are usually toxigenic (Dziva *et al.*, 2004). The symptoms of atrophic rhinitis include sneezing, nasal discharge and epistaxis, snout deformation including shortening or twisting of the snout, dark tear staining below the medial canthus of the eye, pneumonia and decreased growth rates (Adlam & Rutter, 1989). The distorted snout is chronic and does not cause mortality of the infected pigs, but reduces growth rates. Severe turbinate atrophy is frequently caused by toxigenic type D strains. The toxigenic strains produced cytotoxic and dermonecrotic toxin encoded by the *toxA* gene which has a protein and carbohydrate composition (Adlam & Rutter, 1989; Petersen & Foged, 1989). The *toxA* gene is present only in toxigenic strains of *P. multocida* (Petersen & Foged, 1989). This gene is also present in the toxigenic ovine and caprine strains of *P. multocida* (Shayegh *et al.*, 2010a). Pigs infected by *Bordetella bronchiseptica* are more susceptible to infection by the toxigenic strains of *P. multocida* (Adlam & Rutter, 1989). Dermonecrotic toxins produced by *B. bronchiseptica* can alter conditions in the nasal cavity of pigs and allow colonization

Table 1-1. Different diseases caused by *P. multocida* (Adlam & Rutter, 1989)

Animal host	**Disease**
Birds	- acute fowl cholera - chronic localized infections
Cattle and water buffaloes	- pneumonic pasteurellosis - haemorrhagic septicaemia
Pigs	- pneumonic pasteurellosis - atrophic rhinitis
Sheep	- pneumonic pasteurellosis
Rabbits	- rhinitis or snuffles, sinusitis, pneumonia, conjunctivitis, otitis media, subcutaneous abscesses, chronic bronchopulmonary disease, metritis, genital tract infections and septicaemia
Human	- respiratory tract disease (pneumonia and pleural empyema), infections of the central nervous system (meningitis, cerebral abscess and subdural empyema), neonatal septicaemia, bacteraemia, endocarditis, abdominal infections (spontaneous peritonitis and appendicitis), urogenital infections and soft tissue infections with acute inflammation

by toxigenic *P. multocida* (Adlam & Rutter, 1989). The toxin of *B. bronchiseptica* can induce nasal damage in the nasal tissue and causes turbinate atrophy and pneumonic lesions. However, Brockmeier and Register (2007) found that predisposition to non-toxin producing strains of *B. bronchiseptica* can also support colonization by the toxigenic strains of *P. multocida*. *P. multocida* toxin can stimulate cell proliferation, whereas *B. bronchiseptica* toxin does not (Ohnishi *et al.*, 1998).

1.4.4 Haemorrhagic septicaemia

Haemorrhagic septicaemia is an acute disease caused by serotype B:2 and E:2 isolates in cattle and water buffaloes in Asia (B:2) and Africa (E:2) (Adlam & Rutter, 1989). Dey *et al.* (2007) reported the close association of haemorrhagic septicaemia-causing bovine strains of serotype B:2 with other hosts such as swine and sheep. The infection is mostly by direct contact with carriers and contaminated pasture (Adlam & Rutter, 1989). General symptoms are oedematous swelling in the head and neck, swollen and haemorrhagic lymph nodes and numerous subserous petechial haemorrhages (Adlam & Rutter, 1989). Many of the serotype B isolates associated with haemorrhagic septicaemia produce hyaluronidase and endotoxin (Adlam & Rutter, 1989).

1.4.5 Others

1.4.5.1 Pasteurellosis in rabbits

The most common disease caused by *P. multocida* in rabbits is rhinitis or snuffles (Adlam & Rutter, 1989). Others include sinusitis, pneumonia, conjunctivitis, otitis media, subcutaneous abscesses, chronic bronchopulmonary disease, metritis, genital tract infections and septicaemia (Adlam & Rutter, 1989). Capsular types A and D are mainly responsible for these infections (Adlam & Rutter, 1989). However, Jaglic *et al.* (2006) showed that *P. multocida* serotype F isolates also cause serious infections (fibrinopurulent pleuropneumonia or diffuse haemorrhagic pneumonia) in rabbits.

1.4.5.2 Pasteurellosis in cats and dogs

Cats and dogs can be carriers of *P. multocida* (Adlam & Rutter, 1989). The sites of colonization are the mucosa of the upper respiratory and alimentary tracts of cats, and the nares and oral cavity of dogs (Adlam & Rutter, 1989). Systemic diseases are rare, but there are a few reports of valvular endocarditis, leptomeningitis and a lingual abscess from dogs (Adlam & Rutter, 1989).

1.4.5.3 Pasteurellosis in human

Pasteurellosis in human is usually due to cat and dog bites and subsequent wound infections. *P. multocida* causes respiratory tract disease (pneumonia and pleural empyema), infections of the central nervous system (meningitis, cerebral abscess and subdural empyema), neonatal septicaemia, bacteraemia, endocarditis, abdominal infections (spontaneous peritonitis and appendicitis), urogenital infections and soft tissue infections with acute inflammation (Adlam & Rutter, 1989).

1.5 Mechanisms of pathogenesis

Bacterial pathogenesis is the ability of an organism to establish itself in the host and cause damage (Cowan, 2012). This is a multi-factorial process which depends on host status and bacterial factors. First of all, the bacteria originating outside the host body enter the respiratory tract or other sites of infection. After that, adherence and colonization of a cell surface within the host will begin; this often involves a specific interaction between molecules on the bacterial surface and receptors on the host cells. Once attached, the bacteria begin to multiply and this may lead to invasion in which the bacteria will enter the host cells, spread into the bloodstream and survive the host defences. Finally, the bacteria can cause disease and damage to the hosts by toxin production or induction of inappropriate host responses. Bacterial structures involved in the preceding processes are

called virulence factors. However, in *P. multocida*, mechanisms of pathogenesis and virulence are still not clearly established.

1.5.1 Virulence factors

Virulence or degree of pathogenicity can be caused by single or multiple factors. Known virulence factors of *P. multocida* are described in this section.

1.5.1.1 Capsule

The capsule or capsular polysaccharide is the outermost structure of many bacteria and is known to be an important virulence factor. The capsule may be connected to the bacterial cell surface by covalent interactions to lipid A molecules in Gram-negative bacteria (Russell & Herwald, 2005). Normally, capsule is composed of more than 95% water and repeating single monosaccharide molecules joined by glycosidic linkages (Russell & Herwald, 2005). Different monosaccharide composition, order and linkage between these molecules result in capsular diversity (Russell & Herwald, 2005). Functions of the capsule in pathogenicity of many bacteria include resistance to desiccation, adherence, and resistance to nonspecific host immunity via complement-mediated killing and specific host immunity via antibody-mediated killing (Russell & Herwald, 2005). The capsule may be important at the initial stage of colonization of the mucus layer, while the expression of capsule is reduced at later stages to allow specific interaction of bacterial cell-envelope components to the host cell surface (Russell & Herwald, 2005).

In *P. multocida*, the capsule of capsular type A strains contains hyaluronic acid or hyaluronan which is a polymer of D-glucuronic acid and *N*-acetyl-D-glucosamine, whereas the capsule of capsular type D strains is a polymer of heparin or N-acetylheparosan, and that of capsular type F is a polymer of chondroitin (DeAngelis *et al.*, 2002). There is no chemical linkage found in the capsule of capsular type B strains and its monomers are composed of arabinose, mannose and galactose at a ratio of 0.5:2.0:0.8 (Boyce *et al.*,

2000a). Differences in the capsule compositions in *P. multocida* can be explained at the genetic level. Boyce *et al.* (2000a) compared the capsule biosynthetic loci of the capsular type A and B strains of *P. multocida*. The capsule biosynthetic locus of *P. multocida* capsular type A contains 11 genes: *hexABCD*, *hyaABCDE* and *phyAB*, whereas this locus of capsular type B consists of 15 genes: *cexABCD*, *bcbABCDEFGHI* and *lipAB* (Chung *et al.*, 1998; Boyce *et al.*, 2000a). This is shown in **Figure 1-4**. The *hexABCD* and *cexABCD* gene products are involved in the transport of polysaccharides to the cell surface. The *hyaABCDE* gene products are responsible for the formation of activated sugar monomers and the assembly of the capsular type A polysaccharide, while the *bcbABCDEFGHI* genes are responsible for the formation of capsular type B polysaccharide. The last two genes, *phyAB* and *lipAB*, function in phospholipid substitution of the polysaccharides before translocation. Strains lacking capsules were shown to be less virulent than capsulated strains (Jacques *et al.*, 1993). Recently, Steen *et al.* (2010) compared differences between capsulated and non-capsulated avian strains of *P. multocida* by using whole genome sequencing. The authors found no mutations in the capsule biosynthetic locus of the capsulated and non-capsulated strains, but a single point mutation was observed within the *fis* gene of the non-capsulated strain. The *fis* gene is a growth phase-dependent transcriptional regulator which controls capsule gene expression and other virulence genes of *P. multocida* including a filamentous haemagglutinin-encoding gene *pfhB_2* and a cross-protective surface antigen encoding *PlpE*.

1.5.1.2 Lipopolysaccharide

Lipopolysaccharide (LPS) of Gram-negative bacteria is a glycolipid component of the outer membrane and consists of three major parts: a hydrophobic lipid A, hydrophilic inner and outer core polysaccharide chains and a repeating hydrophilic O-antigen oligosaccharide side chain (Kuhnert & Christensen, 2008). Structure of LPS in Gram-negative bacteria has been described by Kuhnert and Christensen (2008) (**Figure 1-5**).

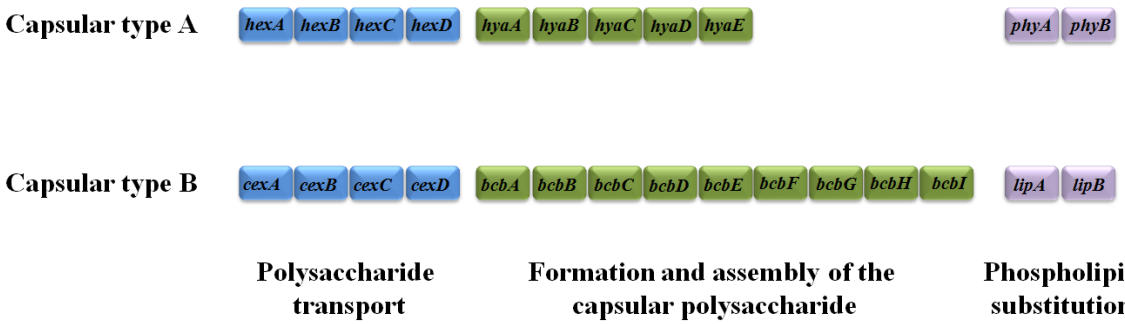


Figure 1-4. Comparison of the capsule biosynthesis loci of *P. multocida* capsular types A and B. The difference is located in the part involved in the formation and assembly of the capsular polysaccharide (Boyce *et al.*, 2000a; 2000b).

Briefly, lipid A is a major component of the outer membrane and has endotoxic properties such as stimulation of cytokines and inflammation. Acylation of lipid A is commonly found. The number and length of the acyl groups attached to lipid A contribute to the host immune recognition and the susceptibility to antimicrobial compounds. The inner core conserved region has a 3-deoxy-D-manno-2-octulosonic acid (Kdo) as a linker and a triheptose unit connected to Kdo. The heptose unit can be extended by the outer core oligosaccharide chains or replaced by phosphoethanolamine (PEtn). However, the structural differences of the outer core region among different bacterial strains can be due to phase variation. The O-antigen of LPS is a repeating oligosaccharide consisting of one to four sugar units but LPS that lacks of the O-antigen is called lipooligosaccharide (LOS).

The LPS of *P. multocida* lacks the O-antigen component (St Michael *et al.*, 2005). A structural study of the LPS of the avian capsular type A strain Pm70 by St Michael *et al.* (2005) showed that it has a triheptose unit linked to a Kdo residue. The first heptose residue is substituted by glucose at the 4- and 6-positions and the second heptose residue is substituted by a phosphoethanolamine residue. However, the glucose residue at the 4-position of the first heptose residue is further substituted by a heptose residue at the 6-position. The extension of this outer core oligosaccharide causes structural variations, resulting in 16 different LPS serotypes of *P. multocida*. Harper *et al.* (2007a) compared the LPS structure of the Pm70 strain to those of the virulent avian strains VP161 and X-73. The authors showed that these virulent strains had phosphocholine residues (PCho) substituted at the terminal galactose residues; these were absent in the Pm70 strain. The PCho residues have roles in adhesion, resistance to antimicrobial peptides and complement-mediated killing. Mutation of the *pcgC* gene, which involves the addition of PCho to LPS, reduced the virulence of the VP161 strain (Harper *et al.*, 2007a). Moreover,

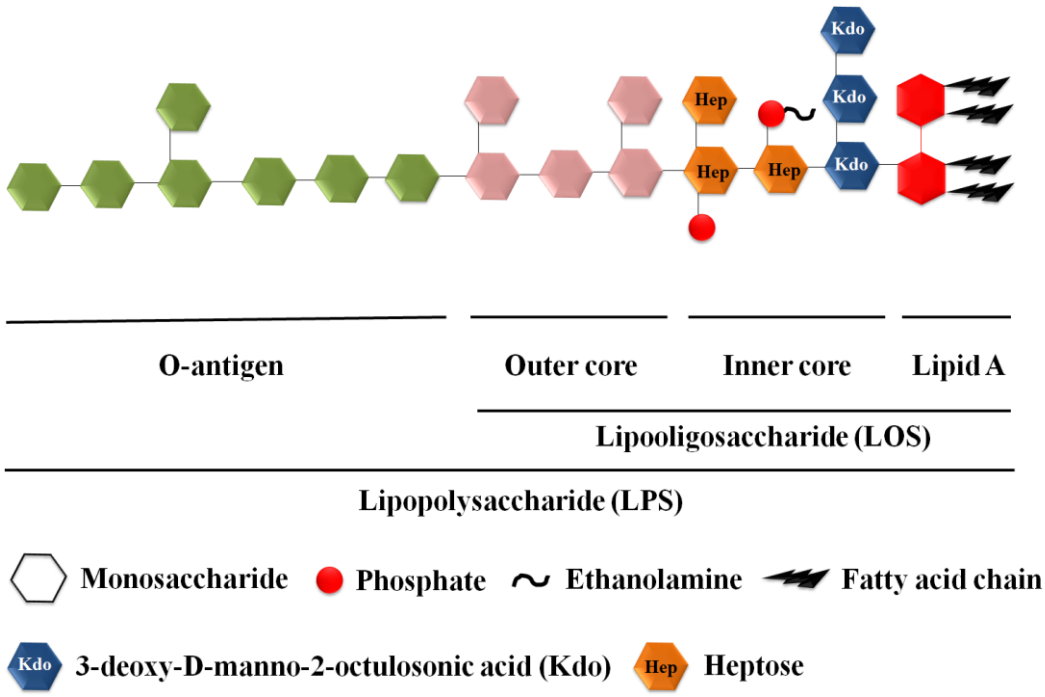


Figure 1-5. Typical structural components of lipopolysaccharide (LPS) which consists of lipid A, inner and outer core oligosaccharides and O-antigen, and lipooligosaccharide (LOS) which consists of lipid A, inner and outer core oligosaccharides, of Gram-negative bacteria (Kuhnert & Christensen, 2008).

Harper *et al.* (2007b) identified two different core inner oligosaccharide structures, one containing a single Kdo residue and the other containing two Kdo residues. The authors found that only one form is required for the virulence of *P. multocida*.

1.5.1.3 Pili and fimbriae

Pili and fimbriae are bacterial adhesive surface appendages contributing to bacterial adherence to host tissues, colonization and biofilm formation (Cowan, 2012). Bundle-forming Type IV fimbriae or pili are highly expressed in capsular type A, B and D strains of *P. multocida* under microaerophilic conditions (Ruffolo *et al.*, 1997). The type IV fimbriae of *P. multocida* are long appendages and consist of the repeating 18-kDa fimbrial subunit PtfA (Adler *et al.*, 1999). The tight adherence (Tad) macromolecular transport system which is required for the assembly of adhesive Flp (Fimbrial low-molecular-weight protein) pili may be responsible for the assembly of the type IV fimbriae of *P. multocida* (Tomich *et al.*, 2007).

The *tad* locus is localized on a mobile genetic island named widespread colonization island (WCI) and is important to biofilm formation and colonization in many Gram-negative bacterial genera including *Actinobacillus*, *Haemophilus*, *Pasteurella*, *Pseudomonas* and *Yersinia* (Tomich *et al.*, 2007). The *tad* locus of *P. multocida* is composed of 14 genes including *flp-1*, *flp-2*, *rcpABC* and *tadABCDEFGVZ* (Tomich *et al.*, 2007). **Figure 1-6** summarises the Tad transport system in *P. multocida*. The *flp-1* gene encodes a major structural component of the Flp pili, while the *flp-2* gene encodes a second pilin. The *rcpA* gene encodes the GspD/PilQ secretin family protein belonging to type II secretion and type IV pilus systems and forms a pore in the outer membrane. The *tadA* gene encodes the GspE ATPase family protein which provides energy for Flp pili assembly. The *tadB* and *tadC* genes encode the GspF/PilC inner membrane proteins which function as a scaffolding unit for Flp pili assembly. The *tadV* gene encodes the GspO/PilD prepilin peptidase. The

tadG gene encodes an inner membrane component of the Flp pili biogenesis apparatus which may anchor the pilus to the cell. The *tadE* and *tadF* genes encode pseudopilins which may function as a piston (pseudopilus) to deliver the substrate from the periplasmic space through the outer membrane. The *tadZ* gene encodes a cytoplasmic protein which may be associated with the inner membrane although its function remains unknown. The *rcpB* gene encodes an outer membrane protein which is believed to stabilize and gate the secretin complex. The *rcpC* gene encodes a predicted outer membrane protein which may be involved in the post-translational modification of the pilin, scaffolding of the outer membrane secretin complex or facilitating the extrusion of the pilus through the secretin pore. The *tadD* gene encodes an outer membrane lipoprotein which functions as a pilotin or pilot protein for proper assembly of the secretin complex. Pili are involved in adhesion and colonization to host cells (Proft & Baker, 2009). Two types of pili occur in porcine isolates of *P. multocida* associated with atrophic rhinitis: rigid pili which lie along the side of the outer membrane and curly pili which are difficult to visualize (Isaacson & Trigo, 1995). Pili were also detected in capsulated and non-capsulated avian *P. multocida* strains P-1059I , P-1059B and P1059G on blood agar, on dextrose-starch agar and in broth (Rebers *et al.*, 1988). Those strains grown on blood agar were heavily piliated.

1.5.1.4 Outer membrane proteins (OMPs)

Several outer membrane proteins of *P. multocida* are involved in virulence such as adherence, iron-uptake, drug efflux and *in vivo* survival. In this section, certain OMPs are given as examples and more details will be provided in **section 1.6.1.3**. The highly abundant OmpA is a surface-exposed and immunogenic protein which functions in adherence (Dabo *et al.*, 2003), invasion (Prasadarao *et al.*, 1996), biofilm formation (Ma & Wood, 2009a), immune evasion (March *et al.*, 2011) and interaction with bacteriophages (Morona *et al.*, 1985). The outer membrane porin OmpH is another highly abundant

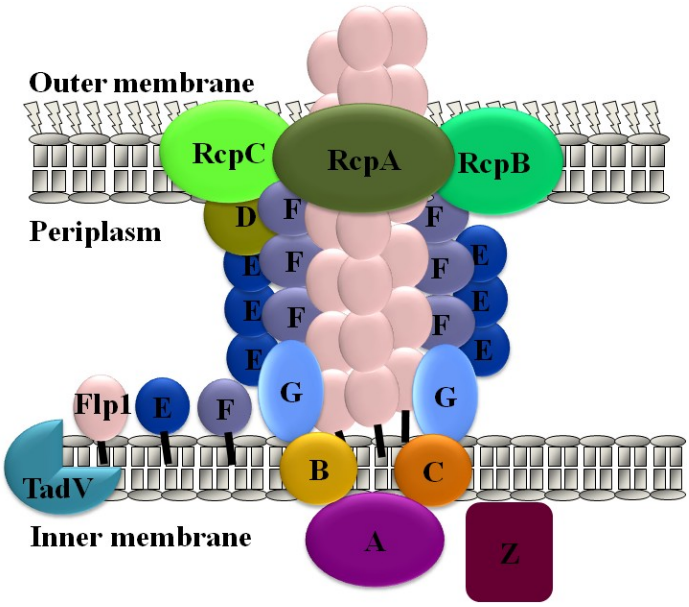


Figure 1-6. Modified model (Tomich *et al.*, 2007) representing the Tad transport system of the Flp pilus of *P. multocida* which contains 14 proteins: Flp1, Flp2, RcpABC and TadABCDEFGVZ. RcpABC forms a pore in the outer membrane. Flp2 is similar to Flp1 and is not shown in the figure.

protein which is surface-exposed (Hatfaludi *et al.*, 2010). Oma87 or Omp87 is a surface-exposed OMP and an antibody against this protein showed homologous protection of the same avian strain (Adler *et al.*, 1999). Interestingly, a surface-exposed and highly immunogenic lipoprotein, PlpE, of *P. multocida* shows high levels of cross protective immunity (Hatfaludi *et al.*, 2010). Certain transport proteins include a multidrug efflux protein TolC involved in multidrug resistance and protein export, and *P. multocida* also express a number of outer membrane iron receptors, e.g. a transferrin receptor TbpA, haemoglobin receptors HemR, HgbA and HgbB, and a haem receptor HasR, which allow them to survive *in vivo* (Hatfaludi *et al.*, 2010). For nutrient acquisition, sialidases NanH and NanB scavenge sialic acid from the host (Hatfaludi *et al.*, 2010). Some OMPs mediate the adherence of *P. multocida* to the host cell surface, for example the TadD and RcpAB proteins of the Tad transport system (Hatfaludi *et al.*, 2010).

1.5.1.5 Secreted proteins

Secreted proteins are proteins that are transported to extracellular environments. Certain of these proteins can be completely released from bacterial cells, whereas others remain anchored to the outer membrane. Protein secretion is a basic function of bacteria and the secreted proteins have various roles e.g. nutrient acquisition, toxins, adherence, colonization, motility, biogenesis of the outer membrane and capsule and intercellular communication. Protein secretion systems in Gram negative bacteria can be generally classified into two groups: Sec-dependent and Sec-independent pathways **(Figure 1-7)** (Gerlach & Hensel, 2007).

The Sec machinery **(Figure 1-7)** contains two inner membrane protein complexes: SecYEG and SecDFYajC (Gerlach & Hensel, 2007). The newly translated preproteins are recognized by an accessory protein SecA together with a chaperone SecB, and delivered to the SecYEG pore complex. The SecDFYajC complex assists the translocation of the

preproteins to the periplasmic space. The signal peptides will be cleaved and the mature proteins will be released into the periplasmic space. Subsequently, the proteins can be secreted through the outer membrane via Type II and V secretion systems. The Type II secretion system **(Figure 1-7)** involves the secretion of enzymes and toxins across the outer membrane. In *Klebsiella oxytoca*, this system is used for the secretion of pullulanase PulA. PulA is translocated through the inner membrane PulFGHIJKLMN protein complex with the energy supplied by PulE. After that, PulA is secreted through the outer membrane protein PulD together with a lipoprotein PulS. The Type V secretion system **(Figure 1-7)** includes autotransporter secretion and two-partner secretion systems. The autotransporter protein consists of three domains: the N-terminal signal sequence, the passenger domain which is translocated across the outer membrane by the β-barrel-forming C-terminal translocation domain. On the other hand, the two-partner secretion system is composed of two proteins which are the passenger protein and the β-barrel outer membrane transporter.

The Sec-independent pathways **(Figure 1-7)** include Type I, III and IV secretion systems (Gerlach & Hensel, 2007). The Type I secretion system **(Figure 1-7)** involves the direct secretion of various substrates from the cytoplasm to the extracellular environment. In uropathogenic *E. coli*, this secretion system consists of an inner membrane ABC transporter HlyB, a membrane fusion protein HlyD and an outer membrane protein TolC, and is used for the secretion of α-hemolysin HlyA. The Type III secretion system **(Figure 1-7)** is a large protein complex which translocates proteins through the inner membrane, the periplasmic space, the outer membrane, the extracellular space and the host cell membrane in to the cytoplasm of the host cell. This complex is composed of a basal body, an inner membrane ring complex, a periplasmic inner rod, an outer membrane

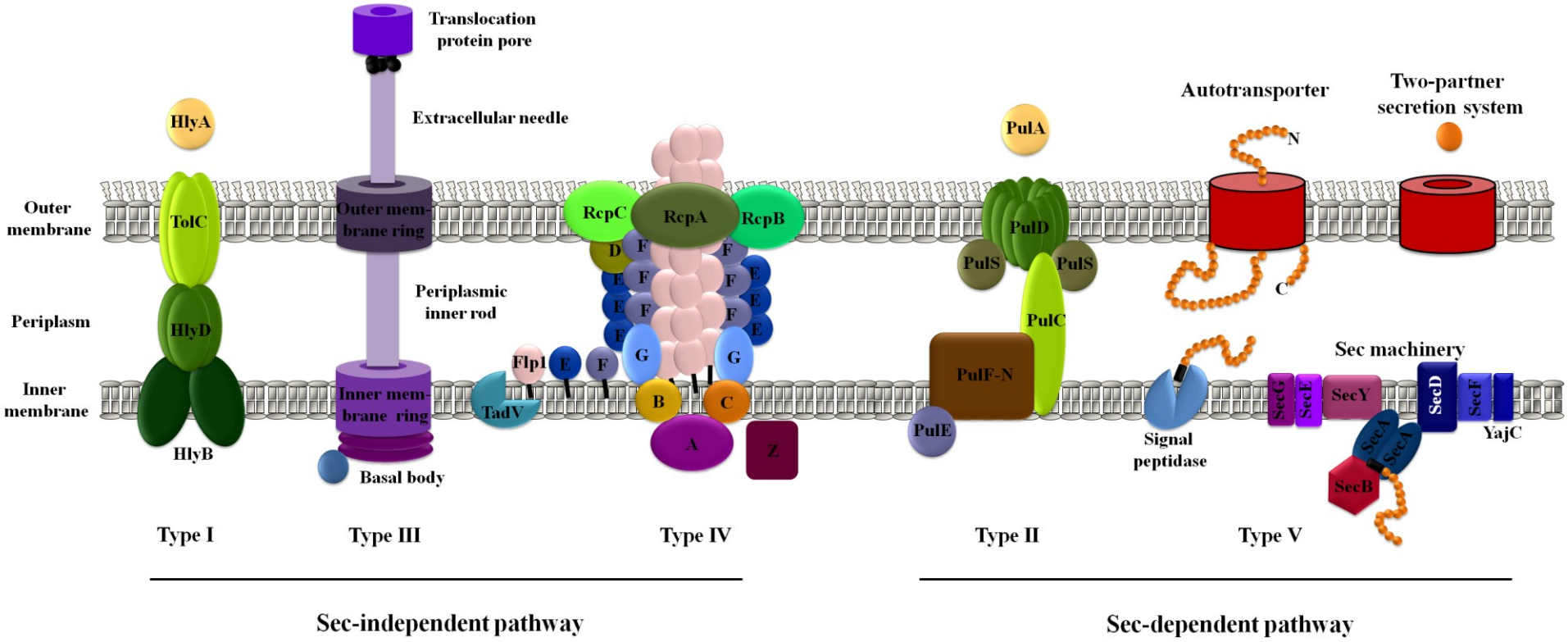


Figure 1-7. Protein secretion systems in Gram-negative bacteria can be classified into Sec-independent and Sec-dependent pathways. The Sec-independent pathways include Type I (*E. coli*), Type III (*Salmonella enterica*) and Type IV (*P. multocida*) secretion systems, whereas the Sec-dependent pathways include Type II (*K. oxytoca*) and Type V secretion systems (Gerlach & Hensel, 2007). Details about each secretion system are described in the text.

secretion ring complex, an extracellular needle protein complex and a translocation protein pore which is formed in the cytoplasmic membrane of the host cell (Gerlach & Hensel, 2007). The Type IV secretion system has been described in **section 1.5.1.3**.

Additionally, another secretion system, named the twin-arginine (Tat) translocation pathway **(Figure 1-8)**, is involved in the secretion of folded or cofactor-containing proteins (Natale *et al.*, 2008). This secretion system has been described by Natale *et al.* (2008). The Tat system consists of two to three inner membrane proteins, TatABC, which translocate folded proteins such as amidase AmiC or cofactor-containing proteins such as oxido-reductases across the inner membrane. Proteins which are translocated by this secretion system are folded within the cytoplasm by the general chaperones such as chaperones DnaK and SlyD. These proteins contain twin-arginine motifs at the interface between the N-terminal positively charged region and the hydrophobic region of their signal sequences. This twin-arginine motif is described as Z-R-R-x-Φ-Φ, where Z is for any amino acid, R for arginine, x for any amino acid and Φ for hydrophobic amino acids (Natale *et al.*, 2008). The folded proteins bind to the TatBC complex via the Tat signal peptides and the TatA protein is recruited to form a translocase pore complex for the translocation of the folded proteins (Natale *et al.*, 2008).

P. multocida uses different secretion systems (including Type I, IV and V secretion systems) to transport various proteins across the inner and outer membranes including a dermonecrotoxin (DNT) or *P. multocida* toxin (PMT), adhesins and proteases (Kuhnert & Christensen, 2008). The dermonecrotoxin is encoded by the *toxA* gene and is associated with swine atrophic rhinitis (Buys *et al.*, 1990). This toxin is a mitogen or a cyclomodulin which promotes proliferation of various cell types and can induce osteoclastic bone resorption at the nasal turbinates and inflammation of the nasal mucosa (Nougayrede *et al.*, 2005; Hildebrand *et al.*, 2011). However, the secretion system responsible for the

dermonecrotoxin transport has not been elucidated. The Flp or type IV pili (type IV secretion system) and trimeric autotransporters (type V secretion system) are adhesins found in *P. multocida* (Kuhnert & Christensen, 2008). These proteins are synthesized and secreted by *P. multocida* to mediate adherence to the host tissues. Certain autotransporters belonging to the type V secretion system have enzymatic functions such as a sialidase NanB which is involved in acquisition of sialic acid from the environment (Hatfaludi *et al.*, 2010). *P. multocida* also secretes neutral metalloproteases which degrade immunoglobulin (Negrete-Abascal *et al.*, 1999).

1.5.2 Host-pathogen interactions

The interactions between bacteria and host are complex. The host organisms can be considered as bacterial environments and bacteria have evolved mechanisms which allow them to respond to different host niches. This section will focus on bacterial adherence, host specificity and response to host environments.

1.5.2.1 Adherence and colonization

Generally, bacterial adherence happens in two steps (Ofek *et al.*, 2003). First, nonspecific adherence occurs; this is a reversible interaction between hydrophobic molecules. Second, firm adherence of bacteria occurs by nonhydrophobic interactions. This happens after the successful completion of the first step by specific interactions of bacterial adhesins and complementary receptors on the host cell surfaces. There are three types of interactions between adhesins and receptors; these include lectin-carbohydrate interactions, protein-protein interactions, and hydrophobic molecule-protein interactions (Ofek *et al.*, 2003). The first type is the binding of lectin with carbohydrate structures. Bacterial lectins can be parts of fimbriae or outer membrane components and bind to glycolipids on the host surface. If the lectins are present on the host surface, they bind to bacterial capsular

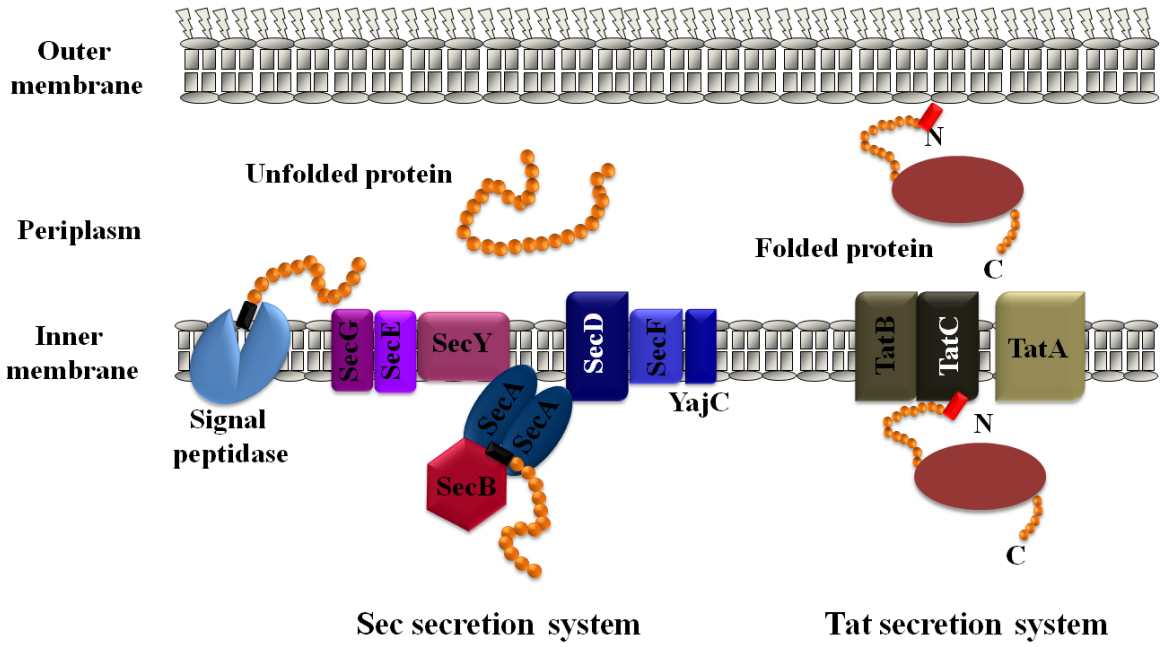


Figure 1-8. Comparison of the Sec and Tat secretion systems (Natale *et al.*, 2008). The Sec secretion system transports unfolded proteins across the inner membrane, while the Tat secretion system transports folded proteins.

polysaccharide or LPS. The second type is the interaction between bacterial proteins such as fibronectin-binding proteins and host proteins on the cell surface including fibronectin and other extracellular matrix components such as collagen, elastin, fibrinogen, laminin and vitronectin. The third type is the interaction of hydrophobic moieties of proteins with lipids. However, the interaction between bacterial adhesins and host cell receptors can be complex and bacteria can have multiple adhesins in order to confront the numerous receptors in different environments.

In *P. multocida*, a comparative adherence study by Glorioso *et al.* (1982) showed that all strains of capsular type A and one strain of capsular type D were highly adhesive to the mucosal epithelium of the nasopharynges of rabbits compared to other strains of capsular types B, D and E. The authors also showed that fimbriae were produced in the highly adhesive strains and that capsule removal increased adhesion. Conversely, an adhesion study of avian strains of *P. multocida* to chicken embryo fibroblast cells showed that the adhesive properties of the capsulated strains were higher than the non-capsulated strains (Borrathybay *et al.*, 2003). These authors also revealed that a 39 kDa protein in the capsule may have a role in adherence to the fibroblast cells. Another study showed the involvement of OmpA of *P. multocida* capsular type A in adherence to the host cells by interacting with heparin and fibronectin (Dabo *et al.*, 2003). Recently, Mullen *et al.* (2008a; 2008b) characterized a novel fibronectin-binding protein ComE1 (PM1665) from a pig strain of *P. multocida*. This protein binds to fibronectin via two helix-hairpin-helix motifs. These are also able to bind to DNA and are involved in natural transformation. Taken together, the adherence of *P. multocida* is a complex process which involves a number of adhesins.

1.5.2.2 Host specificity

Host specificity of bacteria may evolve by positive selection imposed by the host which contributes to different host specificities of different strains of bacteria. The correlation between differential protein expression and host specificity of bacteria has been described in many bacterial species (Ewers *et al.*, 2006; Eswarappa *et al.*, 2008; Lawrence *et al.*, 2010). A molecular evolution and phylogenetic study in different serovars of *Salmonella enterica* which cause diseases in avian and mammalian hosts identified virulence genes conferring host specificity to different serovars of *S. enterica* (Eswarappa *et al.*, 2008). The genes that were differentially evolved in different serovars were components of the type III secretion system. A translocation channel or translocon protein encoded by the gene *sipD*, which translocates secreted proteins into the host cells, is differentially evolved in human serovars compared to the other serovars. Translocons encoding genes *sseC* and *sseD* are essential for the proliferation of these bacteria in the host. The *sseF* gene encodes an effector protein that is secreted into the cytoplasm of the host cell and this protein was conserved in human serovars and serovars that can infect multiple hosts. These proteins were localized in the gene clusters called Salmonella pathogenicity islands which may be acquired by horizontal gene transfer. The authors proposed that these genes were differentially evolved for the host specificity of different serovars. A comparative genomic study of three *M. haemolytica* isolates including a bovine isolate of serotype A1 and two isolates of serotype A2 from bovine and ovine, showed a number of genes specific to each isolate such as a novel effector protein of the type III secretion system (Lawrence *et al.*, 2010). These authors also identified single nucleotide polymorphism (SNP) variation between isolates in a gene encoding O-sialoendoglycopeptidase which is an important enzyme in the host-specific colonization process. Another study in *P. multocida* showed evidences of genes conferring host specificity (Ewers *et al.*, 2006). These authors found the presence of the transferrin binding protein encoding gene *tbpA* was restricted to ruminant strains. The dermonecrotic toxin encoding gene *toxA* was detected in swine,

small ruminants, cattle and poultry strains, but there was a strong association to the disease status in swine.

1.5.2.3 Response to host environments

Colonization of bacteria within the host avoids the competition with other bacteria and allows them to access environmental niches essential for bacterial growth (Storz & Hengge, 2011). These niches can be varied in different animal host species and in different individual animals. However, the host environment is not safe for bacterial survival and the bacteria need to counteract the host defense mechanisms. These mechanisms cause various stress conditions on the colonizing bacteria. Examples of these stress conditions include competition with resident bacteria, nutrient starvation, iron limitation and membrane damage by antimicrobial peptides and complement. The resident bacteria assist their hosts by producing antimicrobial factors which prevent colonization by external bacteria. These resident bacteria can be reduced by the use of antibiotics or by the disease status of the host. Overcoming nutrient restriction and iron limitation within the host are essential for the persistence of the bacteria inside the host. Bacterial strains from different animal hosts can differentially express iron receptors and iron acquisition systems due to the types of available host iron-binding compounds and specificity between the iron-binding compounds and the receptors. Veken *et al.* (1996) examined the ability of *P. multocida* strains of serotype B and E associated with haemorrhagic septicaemia to bind and use transferrin, lactoferrin and haemoglobin. The authors found that all strains could utilize haemoglobin but not lactoferrin. Some strains of serotype B could use transferrin as their iron source. Another stress condition is the membrane damage caused by antimicrobial peptides and complement. The antimicrobial peptides are short, positively charged peptides secreted by the host cells which can damage the outer membrane of Gram-negative bacteria, affect nucleic acid and protein synthesis, and disrupt enzymes. Details of these stress conditions are further explained in **Chapter 4**.

Although bacteria have to confront different stress conditions in the host environment, stress response mechanisms allow the survival of the bacteria in their specific hosts (Storz & Hengge, 2011). The formation of biofilms can protect bacteria against phagocytosis, antibody activity and other immune defensive mechanisms. Bacteria can also modify their surface antigens such as the composition of LPS and the expression of additional surface structures including slimes and capsules to avoid the host immune responses. Other mechanisms, such as toxins, protein secretion systems and adhesins, are used against the host and to inactivate the defense mechanisms.

1.6 The Gram-negative bacterial cell envelope

1.6.1 Cell envelope structure

Similar to other Gram-negative species, *P. multocida* has a simple intracellular structure including nucleoid (genetic material) and ribosomes (protein synthesis machinery). Generally, the outer surface of Gram-negative bacteria has a multi-laminar structure as described below (Costerton *et al.*, 1974).

1.6.1.1 Inner membrane

From the innermost part of the cell surface, the cytoplasmic or inner membrane has a symmetrical phospholipid bilayer structure. The inner membrane contains numerous α-helical transmembrane proteins and membrane-associated proteins involved in the transport of nutrients, ions, and waste products and energy generation. The inner membrane harbours enzymes involved in the synthesis of the basic units of peptidoglycan, lipopolysaccharides and phospholipids (Costerton *et al.*, 1974; Adlam & Rutter 1989). Facey & Kuhn (2010) described the biosynthesis of the inner membrane proteins. The process begins with the translation of proteins at the ribosomes in the cytosol. Because the membrane proteins are hydrophobic, the cytoplasmic chaperones including SecB and signal recognition particles (SRP) will interact with the newly synthesized proteins to

prevent premature folding. The cytoplasmic proteins are properly folded by the ATP-dependent DnaK and GroEL chaperone system.

Proteins which are secreted through the inner membrane **(Figure 1-9)**, such as periplasmic, outer membrane and extracellular proteins, contain a cleavable signal peptide at the N-terminal. This signal peptide has an average length of 20 amino acids with a tripartite structure containing a positively charged N-terminal (n-region), a hydrophobic core (h-region) and a polar C-terminal (c-region) (Natale *et al.*, 2008). The c-region is recognized by signal peptidase I for extracellular proteins and signal peptidase II for OMPs. This c-region also contains a lipoprotein box which is essential for the modification of lipoproteins. The h- and n-regions are recognized by the chaperones.

The integral inner membrane proteins do not have the signal peptide but the signal for integration into the inner membrane is located in the hydrophobic transmembrane segments **(Figure 1-9)**. These hydrophobic segments of the unfolded inner membrane proteins are recognized by the chaperone SRP, whereas the unfolded periplasmic, outer membrane and extracellular proteins are bound to the chaperone protein SecB. The SecB-bound and SRP-bound unfolded proteins are delivered to a peripheral component of the protein secretion complex SecA and a membrane receptor FtsY, respectively. After that, both SecA and FtsY deliver the unfolded proteins to the Sec translocation protein complex. The core of this translocation protein complex consists of the integral inner membrane proteins SecY, SecE and SecG. By ATP hydrolysis, the unfolded proteins are then translocated through the pore of this complex or inserted into the inner membrane by

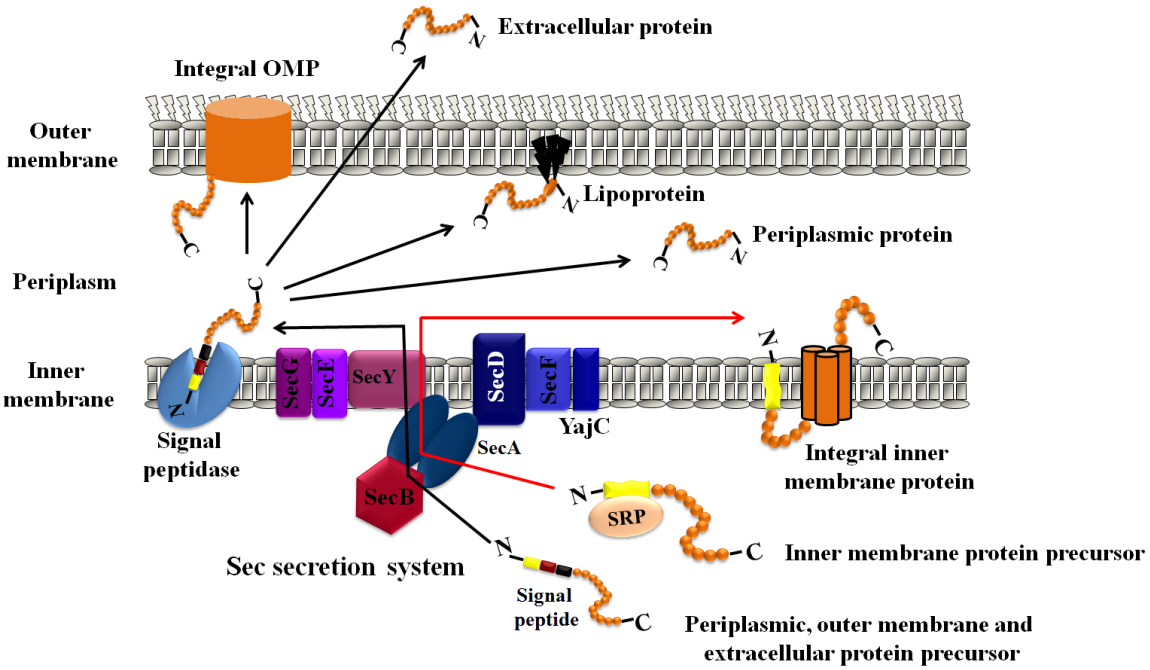


Figure 1-9. The role of the Sec secretion system in the transport and insertion of various types of proteins (Natale *et al.*, 2008). Once translocated across the inner membrane, signal peptides of the periplasmic, outer membrane, and extracellular proteins are processed before transporting to their destinations. Lipoproteins are lipidified prior to cleavage of their signal peptides which consist of three parts: n-region (yellow), h-region (red) and c-region (black). The inner membrane proteins with an inner membrane insertion signal peptide (yellow) are inserted into the inner membrane by the Sec secretion system.

lateral opening of this complex. The additional SecDFYajC protein complex helps the insertion of proteins into the inner membrane. Alternatively, small unfolded inner membrane proteins can be inserted into the membrane by the protein YidC. Once inserted into the inner membrane, the proteins are folded and assembled into the protein complexes such as the Sec complex, the respiratory chain complex, the photosynthesis complex, the F_1F_o ATP synthase complex, the ion and nutrient-uptake permease complex, the ABC transporters, the phosphoenolpyruvate dependent phosphotransferase system (PTS) transporter complex, the aquaporins and mechanosensitive channels and the flagella motor complex (Facey & Kuhn, 2010).

1.6.1.2 Periplasm

Between the inner and outer membrane is a compartment called the periplasm or periplasmic space **(Figure 1-10)** which has a gel-like composition and contains membrane-derived oligosaccharides (MDO) and various proteins such as hydrolytic enzymes, binding proteins, chaperones and chemoreceptors (Ruiz *et al.*, 2006). The periplasm functions as a trans-shipment region in transport between the interior and exterior of the cell. Typical examples of periplasmic transport are the export of polysaccharides to the cell surface, transport of peptidoglycan polymers, fimbrial and extracellular secreted proteins. Protein folding and trafficking chaperones that build and maintain the cell envelope are located here (Ruiz *et al.*, 2006). Once outer membrane and extracellular proteins are translocated through the inner membrane, their N-terminal signal sequence is cleaved and the proteins are folded and assembled in the periplasm (Miot & Betton, 2004). Two types of protein folding chaperones are found in the periplasm: disulphide isomerases (Dsb) which catalyze the formation of disulphide bonds and peptidyl-prolyl isomerases (PPIases such as PpiA, FkpA and SurA) which catalyze the *cis-trans* isomerisation of peptide bonds (Miot & Betton, 2004). Misfolded proteins are degraded by periplasmic proteases such as DegP (Miot & Betton, 2004).

The peptidoglycan or murein layer embedded in the periplasm is a heteropolymer of glycan chains consisting of *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramoyl-peptides (MurNAc) crosslinked by short peptide chains (Bouhss *et al.*, 2008). The peptidoglycan has roles in preventing osmotic stress, maintenance of bacterial cell shape and anchoring other cell envelope components including proteins, polysaccharides, the outer membrane of Gram-negative bacteria and the capsule of Gram-positive bacteria (Bouhss *et al.*, 2008). This layer is thicker in Gram-positive bacteria than it is in Gram-negative bacteria (Cowan, 2012). Biosynthesis of peptidoglycan is a complex process (Bouhss *et al.*, 2008). Briefly, the GlcNAc and MurNAc-pentapeptides precursors are synthesized in the cytoplasm and transferred to the inner membrane receptors. After passage through the membrane, polymerization of the peptidoglycan occurs in the periplasmic space.

Proteins are able to interact with the peptidoglycan by both covalent and non-covalent interactions (Bouhss *et al.*, 2008). These proteins are covalently linked to the peptidoglycan by sortases in Gram-positive bacteria and by L,D-transpeptidases in Gram-negative bacteria (Bouhss *et al.*, 2008). In Gram-negative bacteria, an outer membrane murein lipoprotein or Braun lipoprotein (Lpp) is covalently interacted with the peptidoglycan by its C-terminal residues (Bouhss *et al.*, 2008). The Lpp protein interacts with the Tol/Pal protein complex, consisting of an inner membrane TolA, periplasmic proteins TolB and YbgF and an outer membrane lipoprotein Pal, to maintain the integrity of the cell envelope (Bouhss *et al.*, 2008). In *P. multocida*, other OMPs apart from Lpp are associated with the peptidoglycan including OmpA, and lipoproteins PM0554 and PM0966 (Hatfaludi *et al.*, 2010).

1.6.1.3 Outer membrane

The outer membrane is the second membrane of the Gram-negative bacterial cell envelope. Atypical Gram-positive bacteria named mycolata such as *Mycobacterium smegmatis* have outer membranes containing a thick layer of mycolic acid attached to peptidoglycan (Pajón et al., 2006). The outer membrane functions as a molecular sieve or barrier which limits the size of molecules that can pass through (Ruiz *et al.*, 2006). This has been confirmed by Jaroslawski *et al.* (2009) who showed a molecular sieve-like structure of the outer membrane of the marine bacterium *Roseobacter denitrificans* and a high abundance of the porin proteins which covered approximately 70% of the membrane surface. The outer membrane is also a platform for the anchoring of surface structures such as pili and fimbriae. The bacteria can also produce polysaccharide capsules to cover the outer membrane as the outermost extracellular protective layer (Ruiz *et al.*, 2006).

1.6.1.3.1 Outer membrane structure

The outer membrane **(Figure 1-10)** is an asymmetric membrane comprising phospholipid on the inner side and LPS on the outer surface with certain proteins covalently linked to the peptidoglycan (Ruiz *et al.*, 2006). The overall phospholipid composition of the outer membrane is enriched in saturated fatty acids and phosphatidylethanolamine (Ruiz *et al.*, 2006). The structure and composition of LPS are described in **section 1.5.1.2**. The LPS layer appears gel-like because of the strong lateral interactions between the LPS molecules (Ruiz *et al.*, 2006). In addition to lipids and LPS, there are various proteins localized at the outer membrane. These proteins are generally classified into two groups: integral and peripheral outer membrane proteins. Most of the integral OMPs are transmembrane β-barrel proteins such as porin proteins. The β-barrel structure is also found in mitochondrial and chloroplast OMPs in eukaryotes supporting the endosymbiont theory. About 90% of the peripheral OMPs such as lipoproteins are anchored to the inner side of the membrane and are covalently linked to the peptidoglycan or non-covalently associated with the

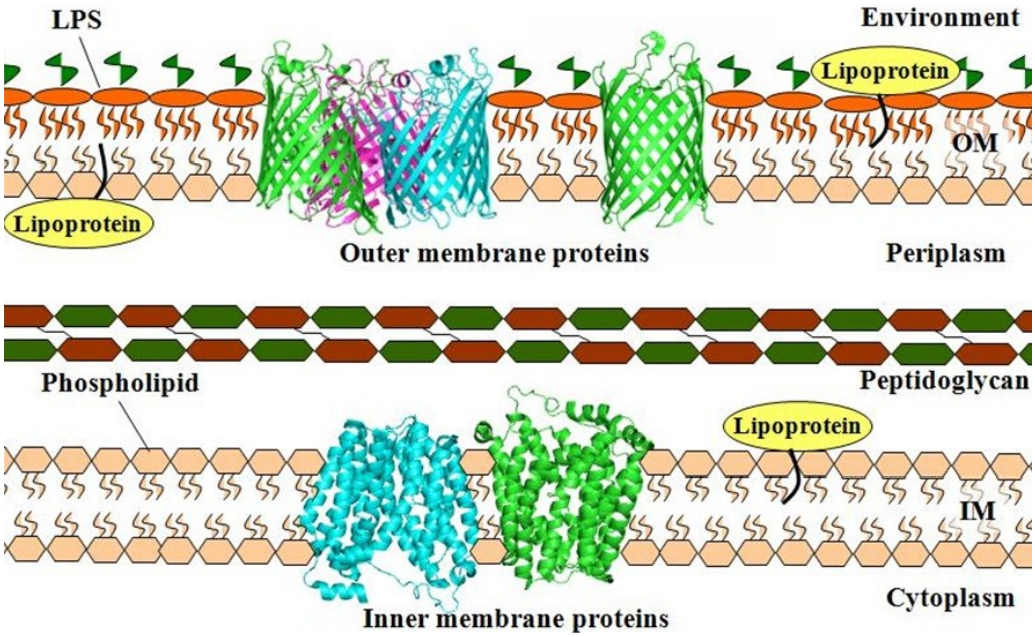


Figure 1-10. Structure of the Gram-negative cell envelope consisting of inner and outer membranes (Ruiz *et al.*, 2006). Transmembrane α-helical proteins mostly occupy the inner membrane, whereas transmembrane β-barrel proteins are common in the outer membrane. Lipoproteins are associateded with both the inner or outer membranes.

peptidoglycan, whereas certain peripheral proteins are surface exposed (Ruiz *et al.*, 2006). OMPs have a variety of roles including biogenesis and integrity of the outer membrane, transport function, adherence and enzymatic activity(Hatfaludi *et al.*, 2010).

1.6.1.3.2 Outer membrane biogenesis

The outer membrane integral proteins and lipoproteins are synthesized in the cytoplasm, whereas the phospholipids and LPS are synthesized at the cytoplasmic side of the inner membrane **(Figure 1-11)** (Ruiz *et al.*, 2006). The OMPs containing the N-terminal signal sequence are translocated across the inner membrane Sec machinery into the periplasmic space (Ruiz *et al.*, 2006). Once translocated, only the outer membrane lipoproteins are lipidified. The N-terminal signal peptides of both integral OMPs and peripheral outer membrane lipoproteins are cleaved by the signal peptidases at the outer side of the inner membrane **(Figure 1-11)** (Ruiz *et al.*, 2006). The unfolded OMPs possibly bind to periplasmic chaperones and protein-folding factors (Skp, SurA, and DegP) before insertion into the outer membrane **(Figure 1-11)** (Ruiz *et al.*, 2006; Knowles *et al.*, 2009). The Skp periplasmic protein can bind many unfolded and newly synthesized OMPs to prevent aggregation and the periplasmic peptidyl-prolyl isomerase SurA helps folding of OMPs and also functions as a chaperone (Ruiz *et al.*, 2006). DegP is a protease which has chaperone activity (Knowles *et al.*, 2009). SurA might be responsible for the assembly of most OMPs, whereas Skp and DegP help those which are misfolded by the first pathway (Knowles *et al.*, 2009). The β-barrel assembly machinery (BAM) consisting of an integral membrane protein Omp85/YaeT/BamA and four accessory lipoproteins, YfgL/BamB, NlpB/BamC, YfiO/ComL/BamD and SmpA/BamE is thought to assist the insertion and folding of the integral β-barrel OMPs into the outer membrane **(Figure 1-11)** (Ruiz *et al.*, 2006; Knowles *et al.*, 2009). BamA has five POTRA (polypeptide transport-associated) domains extending into the periplasmic space which might bind to the unfolded integral OMPs (Knowles *et al.*, 2009). BamCDE interact with BamA via the POTRA domains,

whereas BamB interacts directly with BamA (Knowles *et al.*, 2009). Several models have been proposed for the insertion of OMPs into the outer membrane (Knowles *et al.*, 2009). The first model suggests that the monomeric BAM complex forms a pore for the insertion of OMPs. The second model suggests that oligomerization of BAM complexes occurs and that the OMPs are inserted into the membrane by the central pore. Another model suggests the formation of cage-like multimeric DegP structures which may deliver folded OMPs and insert them into the membrane.

By contrast, after undergoing lipid modification, the outer membrane lipoproteins interact with the ABC transporter complex LolCDE at the inner membrane **(Figure 1-11)** (Ruiz *et al.*, 2006). Although the LolCDE complex is an ABC transporter, it does not transport any substrates because it has only eight transmembrane strands which is less than those of normal ABC transporters (Tokuda & Matsuyama, 2004). The inner membrane lipoproteins do not interact with the LolCDE complex because they contain the Lol avoidance signal (Ruiz *et al.*, 2006). Upon ATP hydrolysis, the outer membrane lipoproteins are released from the inner membrane to the periplasmic chaperone LolA **(Figure 1-11)** (Ruiz *et al.*, 2006). When reaching the outer membrane, the lipoprotein-LolA complex binds to the outer membrane lipoprotein LolB and the lipoprotein is then transferred to LolB and attached to the outer membrane **(Figure 1-11)** (Narita *et al.*, 2004).

For the phospholipids and LPS, after synthesis they are transferred to the outer side of the inner membrane by the ABC transporter MsbA **(Figure 1-11)** (Ruiz *et al.*, 2006). This protein flips LPS and possibly phospholipids from the inner leaflet to the outer leaflet of the inner membrane (Ruiz *et al.*, 2006). Next, the LPS interacts with the ABC transporter LptFGB complex and the inner membrane protein LptC **(Figure 1-11)** (Tokuda, 2009). After that, the LPS is transported to the outer membrane by the periplasmic protein LptA (Tokuda, 2009). Once destined for the outer membrane, LPS molecules are inserted into

the membrane and flipped to the outer side by a transmembrane outer membrane protein Imp/OstA/LptD and an outer membrane lipoprotein LptE **(Figure 1-11)** (Ruiz *et al.*, 2006; Tokuda, 2009; Freinkman *et al.*, 2011). However, the transport of phospholipids to the outer membrane remains unclear.

1.6.2 Outer membrane proteins

The OMPs generally consist of the integral and peripheral proteins. These proteins are essential for bacterial cells and their interactions with the environment and host tissues. Although cell envelope protein-encoding genes account for 20-30% of the bacterial genome, 2-3% of the genome encode the outer membrane proteins (Hatfaludi *et al.*, 2010).

1.6.2.1 The outer membrane proteome

The outer membrane proteome is the entire set of both integral or transmembrane and peripheral or outer membrane-associated proteins present at the outer membrane in a given type of organism at a particular time under defined growth conditions (Brown, 2007). The copy numbers of each protein in the outer membrane proteome vary from rarely found to highly abundant proteins. Examples of highly abundant proteins in the outer membrane of *P. multocida* are the outer membrane porin protein, OmpH, and the structural protein, OmpA (Hatfaludi *et al.*, 2010). However, the size of the outer membrane proteome can vary depending on the bacterial species. The study of the outer membrane proteome mainly by using electrophoresis and mass spectrometry is called outer membrane proteomics and has been used to examine the outer membrane proteomes of various bacterial species such as *Legionella pneumophila* (Khemiri *et al.*, 2008), *Edwardsiella tarda* (Kumar *et al.*, 2009), *Tannerella forsythia* (Veith *et al.*, 2009), *Salmonella enterica* (Chooneea *et al.*, 2010), *Caulobacter crescentus* (Phadke *et al.*, 2001), *P. multocida*

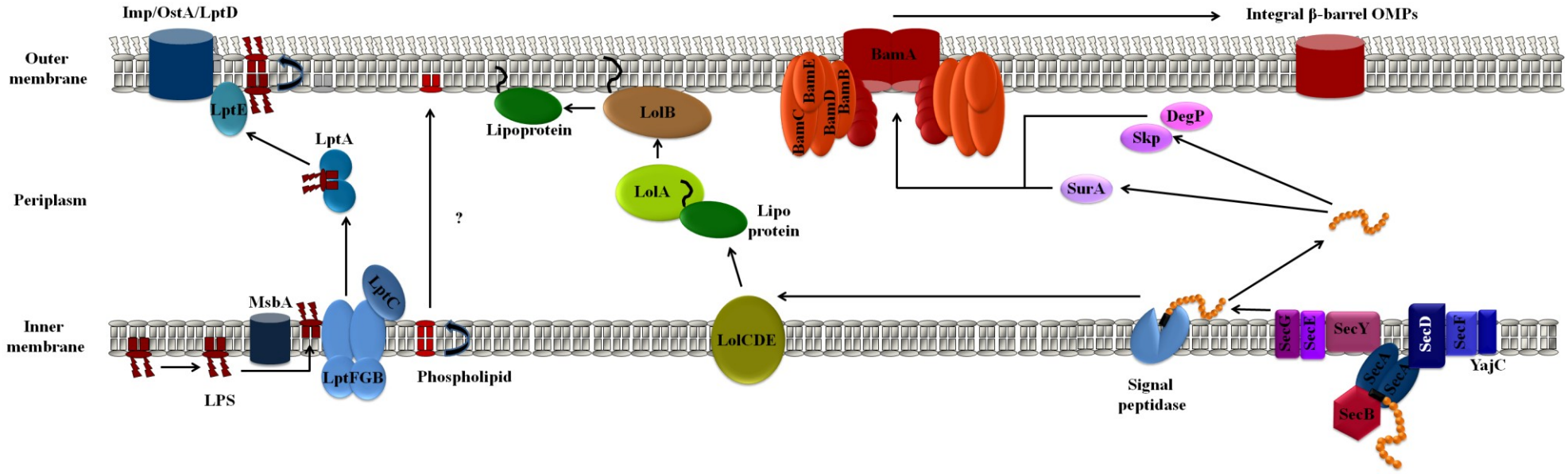


Figure 1-11. Overview of the outer membrane biogenesis in Gram-negative bacteria (Ruiz *et al.*, 2006) which includes biosynthesis of LPS by the MsbA and LptABCDEFG complexes, phospholipids by an unknown mechanism, lipoproteins by the LolABCDE complex and integral β-barrel OMPs by the β-barrel assembly machinery BamABCDE complex together with the periplasmic chaperones and the Sec machinery.

(Boyce *et al.*, 2006), *Prevotella intermedia* (Yu *et al.*, 2007), *Neisseria* species (Abel *et al.*, 2007), *Flavobacterium psychrophilum* (Dumetz *et al.*, 2008), *Dickeya dadantii* (Babujee *et al.*, 2007) and *Escherichia coli* (Molloy *et al.*, 2000).

1.6.2.2 Classification

Outer membrane proteins are generally classified into two groups based on their positions in the outer membrane. Integral or transmembrane proteins are embedded in the outer membrane and mostly have β-barrel structures; peripheral proteins are lipidified and are anchored or associated with either the inner or outer leaflet of the membrane.

1.6.2.2.1 Transmembrane β-barrel proteins

Transmembrane β-barrel proteins are present in the outer membranes of Gram-negative bacteria, Mycobacteria, mitochondria and chloroplasts (Walther *et al.*, 2009). These transmembrane β-barrel OMPs traverse the outer membrane by the antiparallel arrangement of multiple amphipathic β-strands (between 8-22 strands) forming a cylindrical shape (Walther *et al.*, 2009). The antiparallel β-strands have alternate polar and non-polar residues. The polar residues point into the pore forming a polar channel and the non-polar residues interact with the hydrophobic surface of the outer membrane (Valavanis *et al.*, 2006). The transmembrane β-strands are predominantly rich in glycine and the two rings at both ends of the barrel, which interact with the lipid layers, are frequently composed of aromatic amino acids (Tamm *et al.*, 2004). The transmembrane β-barrel OMPs have tight turns on the periplasmic side and variable-sized loops on the extracellular side which contain the functional characteristics of these proteins (Tamm *et al.*, 2001; Tamm *et al.*, 2004). For example, Maruvada & Kim (2011) showed that the extracellular loops of OmpA in *E. coli* contribute to the pathogenesis of neonatal meningitis. Sometimes, one or two loops fold into the pore to form specific-sized pores with moderate substrate specificity such as the maltoporin LamB which is a selective transporter of maltodextrin (Tamm *et al.*, 2004). Many transmembrane β-barrel proteins are monomers,

homo-dimers, such as an outer membrane phospholipase OmpLA, and homo-trimers such as outer membrane porins (Tamm *et al.*, 2004). Transmembrane β-barrel proteins frequently have an N-terminal β-barrel domain and a periplasmic globular C-terminal plug domain such as TonB-dependent receptors (Tamm *et al.*, 2004).

1.6.2.2.2 Lipoproteins

Lipoproteins are anchored to either the inner or outer membranes **(Figure 1-12)**. The inner membrane lipoproteins are anchored to the inner membrane through the hydrophobic residues, whereas the outer membrane lipoproteins are anchored to the outer membrane through the amphipathic β-strands (Tokuda & Matsuyama, 2004). An observation in *E.coli* suggested that most of the lipoproteins are localized at the outer membrane (Tokuda & Matsuyama, 2004). The lipoprotein precursor has a conserved lipoprotein box or lipobox around the signal sequence cleavage site which is processed to form the mature protein **(Figure 1-12)** (Tokuda & Matsuyama, 2004). This lipid modification process occurs at the periplasmic side of the inner membrane **(Figure 1-12)** (Tokuda & Matsuyama, 2004). First, the thioether bond between the N-terminal Cys residue of the protein and diacylglycerol is formed by phosphatidylglycerol/prolipoprotein diacylglycerol transferase (Lgt). Next, the signal peptide sequence is cleaved by prolipoprotein signal peptidase II (LspA). The last step is the aminoacylation of the N-terminal Cys residue by phospholipid/apolipoprotein transacylase (Lnt). This N-terminal lipid is used to anchor the lipoprotein to the periplasmic side of the inner membrane. If the lipoproteins contain Asp at position +2, also called an inner membrane retention signal, these lipoproteins remain anchored to the inner membrane **(Figure 1-12)** (Tokuda & Matsuyama, 2004). Other residues at this position lead to outer membrane localization by the LolAB complex.

1.6.2.2.3 Other proteins

Although transmembrane β-barrel proteins are predominantly found at the outer membrane, a capsular polysaccharide transporter Wza is an unusual outer membrane

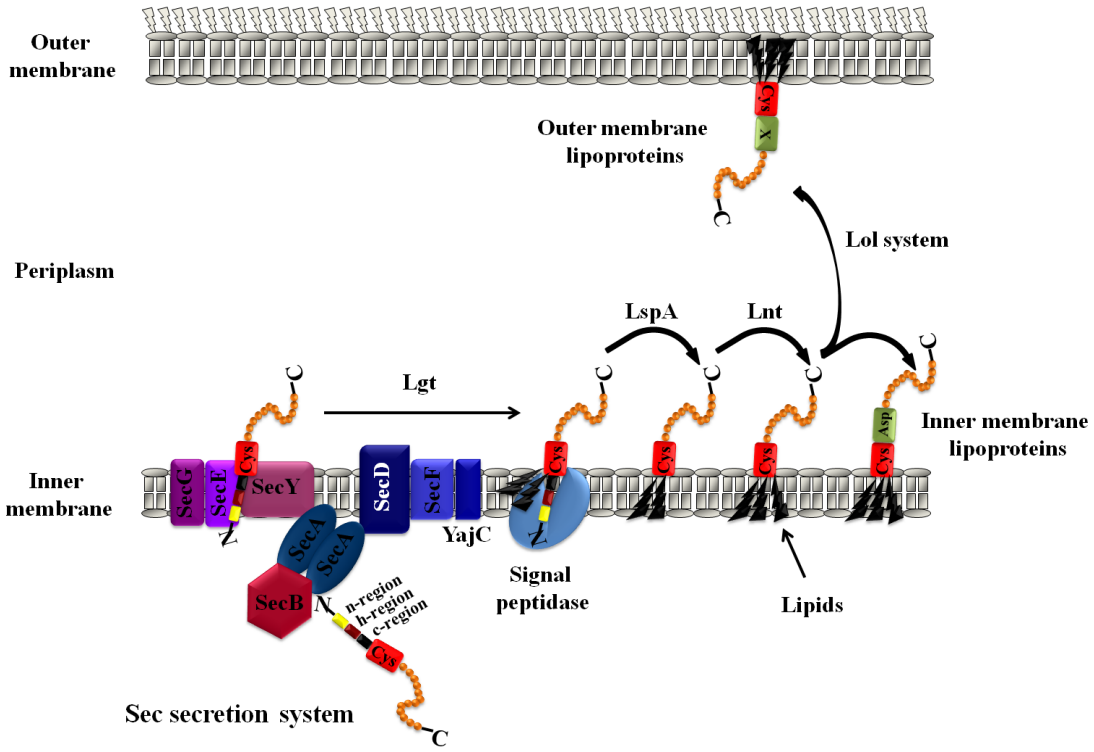


Figure 1-12. Biosynthesis of inner and outer membrane lipoproteins modified from the model in *E. coli* (Tokuda & Matsuyama, 2004). After the lipoproteins are translocated into the periplasmic space by the Sec secretion system, they are further processed by a three-step lipid modification and cleavage of signal peptide. Whether the lipoproteins are destined for the outer membrane or remain at the inner membrane is determined by the amino acid at the position +2; Asp for inner membrane and others for the outer membrane.

lipoprotein which has a transmembrane amphipathic α-helical structure spanning the outer membrane (Ford *et al.*, 2009). This cylindrical-like protein is octameric and each monomer consists of four domains forming four symmetrical rings (Cuthbertson *et al.*, 2009). The first three domains create an internal cavity within the periplasmic space, whereas the forth domain is composed of the amphipathic α-helices which traverse through the outer membrane (Cuthbertson *et al.*, 2009).

The transport of capsular polysaccharide **(Figure 1-13)** begins with the assembly of the capsular polysaccharide precursor at the cytoplasmic side of the inner membrane by the enzyme WbaP (Collins & Derrick, 2007). Next, the capsular polysaccharide precursor is flipped from the cytoplasmic side to the periplasmic side of the inner membrane by the transmembrane inner membrane flippase Wzx (Cuthbertson *et al.*, 2009). After that, polymerization of the capsular polysaccharide occurrs at the periplasmic side by the transmembrane inner membrane polymerase Wzy (Cuthbertson *et al.*, 2009). The capsular polysaccharide is then exported to the surface by the polysaccharide export complex consisting of the periplasmic protein Wzc and the outer membrane lipoprotein Wza which span the cell envelope (Cuthbertson *et al.*, 2009). These polymerization and export processes of the capsular polysaccharide are regulated by the cytoplasmic phosphotyrosine phosphatase Wzb (Collins & Derrick, 2007). Once delivered to the surface, the surface attachment of the capsule may be determined by an outer membrane protein, Wzi (Collins & Derrick, 2007).

1.6.2.3 Functional categories and examples

Many research groups have classified bacterial OMPs into different categories. Tamm *et al.* (2004) generally categorized the OMPs from *E. coli* into six families with respect to their functions including (1) general porins such as OmpC, OmpF and PhoE, (2) passive

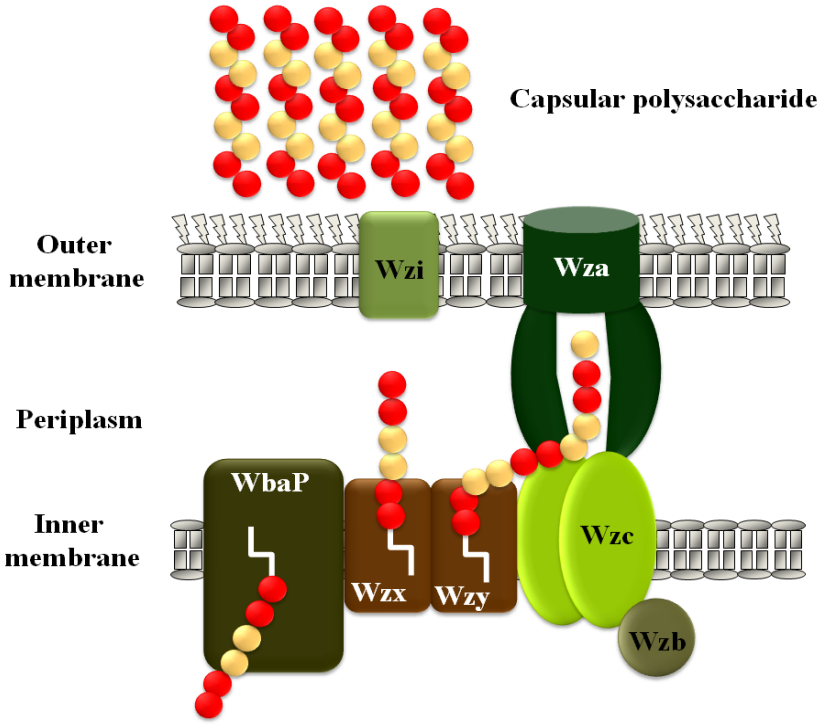


Figure 1-13. The protein complex involved in the biosynthesis and transport of capsular polysaccharide modified from the model in *E. coli* (Collins & Derrick, 2007). The assembly of the capsular polysaccharide precursor (coloured circles) occurrs at the cytoplasmic side of the inner membrane by WbaP. The precursor is flipped from the cytoplasmic side to the periplasmic side of the inner membrane by the flippase Wzx. Polymerization of the capsular polysaccharide occurs at the periplasmic side by polymerase Wzy. The capsular polysaccharide is then exported to the surface by the polysaccharide export complex (Wzc and Wza) and export processes are regulated by the phosphotyrosine phosphatase Wzb.

transporters such as LamB and FadL, (3) active transporters such as the siderophore receptors FepA, FecA and FhuA, (4) enzymes such as phospholipase OmpLA and protease OmpT, (5) defensive proteins such as OmpX and (6) structural proteins such as OmpA. Similarly, in *P. multocida*, Hatfaludi *et al.* (2010) classified the OMPs into six groups based on their functional characteristics including (1) structural proteins, (2) transport proteins, (3) binding proteins, (4) adhesins, (5) protein assembly machines and (6) membrane-associated enzymes.

Structural proteins include the transmembrane β-barrel protein OmpA and the peptidoglycan-associated lipoproteins PCP (PM0554) and Omp16 (PM0966) (Hatfaludi *et al.*, 2010). These proteins maintain the stability of the cell envelope. OmpA links the outer membrane to the peptidoglycan layer by the use of a globular C-terminal domain. OmpA also functions as a virulence factor (it has roles in adherence, as an invasin, in immune invasion and in the formation of biofilms) and as a bacteriophage receptor (Hatfaludi *et al.*, 2010). The structure of OmpA is monomeric with four extracellular loops (Hatfaludi *et al.*, 2010). Heat modifiability of this protein has been described and correlates to the formation of tertiary structures (Tamm *et al.*, 2004). OmpA shows a shift of molecular mass on SDS-PAGE from ~ 30 kDA when the proteins are completely folded to ~ 35 kDa when they are unfolded or incompletely folded after heating at 100ºC. For the lipoproteins, PCP and Omp16 are associated with the peptidoglycan by the C-terminals. The Omp16 protein belongs to the OmpA-like peptidoglycan-associated lipoprotein superfamily (Hatfaludi *et al.*, 2010).

Transport proteins can be divided into two groups: non-specific porins and energy-dependent efflux proteins (Hatfaludi *et al.*, 2010). The non-specific porins allow the passive transport of hydrophilic molecules with poor substrate selectivity and include the trimeric porins OmpC, OmpF and PhoE in *E. coli* and OmpH in *P. multocida* (Hatfaludi *et*

al., 2010). OmpC and OmpF are regulated by osmotic pressure and are involved in bile resistance in *E. coli* (Lin *et al.*, 2002). OmpC has a smaller pore size than OmpF. The phosphoporin PhoE is present under phosphate-limited conditions. In *P. multocida*, OmpH is conserved in many serotypes and the two variable loop regions of this protein may play a role as a serotype-specific epitope (Luo *et al.*, 1999). Sthitmatee *et al.* (2008) identified a 39 kDa (Cp39) protein in *P. multocida* as a capsule-associated adhesin and suggested it was identical to OmpH. This finding has raised a question on the interaction and translocation of proteins between the outer membrane and capsule. The energy-dependent efflux proteins in *P. multocida*, TolC (PM0527) and IbeB (PM1980), are multidrug efflux pumps and components of the Type I secretion system (Hatfaludi *et al.*, 2010). These efflux pumps are involved in the multidrug resistance of bacteria (Hatfaludi *et al.*, 2008; 2010).

Binding proteins include TonB-dependent outer membrane iron receptors **(Figure 1-14)** such as siderophore, transferrin, haem and haemoglobin receptors (Hatfaludi *et al.*, 2010). The iron acquisition system consists of a TonB-dependent outer membrane receptor, a periplasmic binding protein and an inner membrane ABC transporter (Andrews *et al.*, 2003). The energy driving this system is supplied by the inner membrane energy-transducing TonB-ExbB-ExbD system **(Figure 1-14)** (Andrews *et al.*, 2003). The TonB-dependent receptors are transmembrane β-barrel proteins consisting of 22 β-strands with the N-terminal plug or cork domain positioned inside the pores (Andrews *et al.*, 2003; Noinaj *et al.*, 2010). The plug domain is responsible for binding to the extracellular iron compounds and the interaction with the TonB complex via the N-terminal TonB box (Noinaj *et al.*, 2010).

Siderophores are secreted small molecules which function in iron removal from the host iron-containing proteins. The iron-bound siderophores interact with the siderophore

receptors **(Figure 1-14)** such as the iron-siderophore receptor, FepA, and the ferrichrome iron receptor, FhuA, in *E. coli* (Hatfaludi *et al.*, 2010). Once bound to the receptors, the iron-siderophore complex is then transported into the periplasmic space. The siderophores are next transported by the periplasmic binding proteins such as FepB, FhuD and FecB in *E. coli* (Andrews *et al.*, 2003). After that, the siderophores are delivered into the cytoplasm by the ABC transporters such FepBC, FhuBC and FecBC in *E. coli* (Andrews *et al.*, 2003). In *P. multocida*, three proteins with molecular mass of 76 kDa, 84 kDa and 96 kDa were reported as siderophores or multocidin receptors under iron limitation (Hatfaludi *et al.*, 2010).

The transferrin receptor TbpA **(Figure 1-14)** interacts directly with host transferrin (Hatfaludi *et al.*, 2010). TbpA is a transmembrane β-barrel protein which binds to transferrin using its extracellular loops and then removes iron from the transferrin before passing it through the outer membrane. The initial binding process is facilitated by an outer membrane lipoprotein, TbpB. Bovine strains of *P. multocida* bind to transferrin by using only TbpA which has one extracellular loop shorter than other proteins in the TbpA family. This loop was predicted to interact with the TbpB protein. In *N. meningitidis*, the periplasmic binding protein FbpA receives iron from TbpA and transports it to the ABC transporter FbpBC at the inner membrane (Perkins-Balding *et al.*, 2004).

Haemoglobin-binding proteins **(Figure 1-14)** such as HgbA, HgbB and HasR bind to haem, haemoglobin and haemoglobin-haptoglobin as their substrates (Hatfaludi *et al.*, 2010). In *P. multocida*, HgbA and HgbB are haemoglobin-binding proteins (Bosch *et al.*, 2002; Cox *et al.*, 2003). Multiple haemoglobin-binding proteins have been reported in *P. multocida* including PM0400, PM0236, HemR (PM0576), PM0592, PM0741, PM1081, PM1282 and PM0142 (Hatfaludi *et al.*, 2010). The outer membrane receptor HasR **(Figure 1-14)** is a component of the haem acquisition system (Has) which extracts haem

from the secreted haemophore HasA (Wandersman & Delepelaire, 2004; Prado *et al.*, 2005). However, a single haem receptor is able to bind to a variety of haem-containing compounds. This ability was shown in the haem receptor HemR of *Yersinia enterocolitica*; HemR-expressing strains were able to utilize haem, haemoglobin, haptoglobin-haemoglobin, myoglobin, haemopexin and catalase as their iron sources (Bracken *et al.*, 1999). In *P. multocida*, HgbA, PM0400, PM0236, PM0741, PM1081, PM1428 and PM0592 bind to both haem and haemoglobin, whereas PM0576 and PM1282 bind either haem or haemoglonbin (Bosch *et al.*, 2004). However, the periplasmic-binding protein and the ABC transporter responsible for haem/haemoglobin uptake remain unknown.

In addition, certain binding lipoproteins have been reported in *P. multocida* including PlpB/MetQ (PM1730), HemR and PM1578 (Hatfaludi *et al.*, 2010). PlpB functions as a methionine transporter. Although HemR has been known as a haemoglobin receptor, both HemR and PM1578 have putative periplasmic binding domains (Hatfaludi *et al.*, 2010).

Adhesins are involved in host attachment and colonization. *P. multocida* produces a number of outer membrane adhesins including a fibronectin-binding protein ComE1 (PM1665), a trimeric autotransporter adhesin Hsf, filamentous haemagglutinin transporters LspB_1 and LspB_2, and the Tad secretion system (Hatfaludi *et al.*, 2010). ComE1 has dual functions including fibronectin binding and DNA uptake (Mullen *et al.*, 2008). The trimeric autotransporter adhesins have a lollipop-like structure consisting of three domains: head, stalk and anchor domains (Linke *et al.*, 2006). The anchor domain forms a transmembrane β-barrel structure in the outer membrane which functions in the autotransport of the head and stalk domains (Linke *et al.*, 2006a). The stalk domain has an extremely long coiled coil-rich structure and the head domain functions in attachment to the host cells (Linke *et al.*, 2006a). LspB_1 and LspB_2 are components of the two-

partner secretion system which secrete the filamentous haemagglutinin proteins FhaB1 and FhaB2 (Jacob-Dubuisson *et al.*, 2001). The filamentous haemagglutinin proteins have a role in adhesion to the host cells (Hatfaludi *et al.*, 2010). Lastly, the Tad secretion system functions in the production of type IV fimbriae (pili).

Membrane-associated enzymes include a phospholipase, OmpLA, sialidases, NanH and NanB, and a glycerophosphodiester phosphodiesterase, GlpQ (Hatfaludi *et al.*, 2010). OmpLA is a dimeric lipolytic enzyme in which the active site is at the outer part of the interface between two subunits. This phospholipase OmpLA is involved in the hydrolysis of phospholipids in the outer leaflet of the outer membrane (Snijder & Dijkstra, 2000). Sialidases NanH and NanB function in the sequestration of sialic acids from the host sialyl-conjugated glycoproteins or glycolipids (Hatfaludi *et al.*, 2010). The bacteria can use sialic acids as nutrients and can also incorporate the sialic acids into their capsular polysaccharides or LPS to escape from host recognition (Mizan *et al.*, 2000). In *P. multocida*, NanB cleaves both 2-3′ and 2-6′ sialyl lactose, while NanH cleaves only 2-3′ sialyl lactose (Mizan *et al.*, 2000). GlpQ is a non-surface-exposed lipoprotein which removes glycerophosphocholine from the surface of the host cells and converts the glycerophosphocholine to choline for LPS decoration in *H. influenzae* (Hatfaludi *et al.*, 2010).

Protein assembly machinery includes the BAM complex for transmembrane β-barrel protein folding and insertion, the Lol complex for lipoprotein insertion and the LPS insertion proteins Imp/OstA/LptD and LptE. Details of these proteins have been described in **section 1.6.1.3.2**.

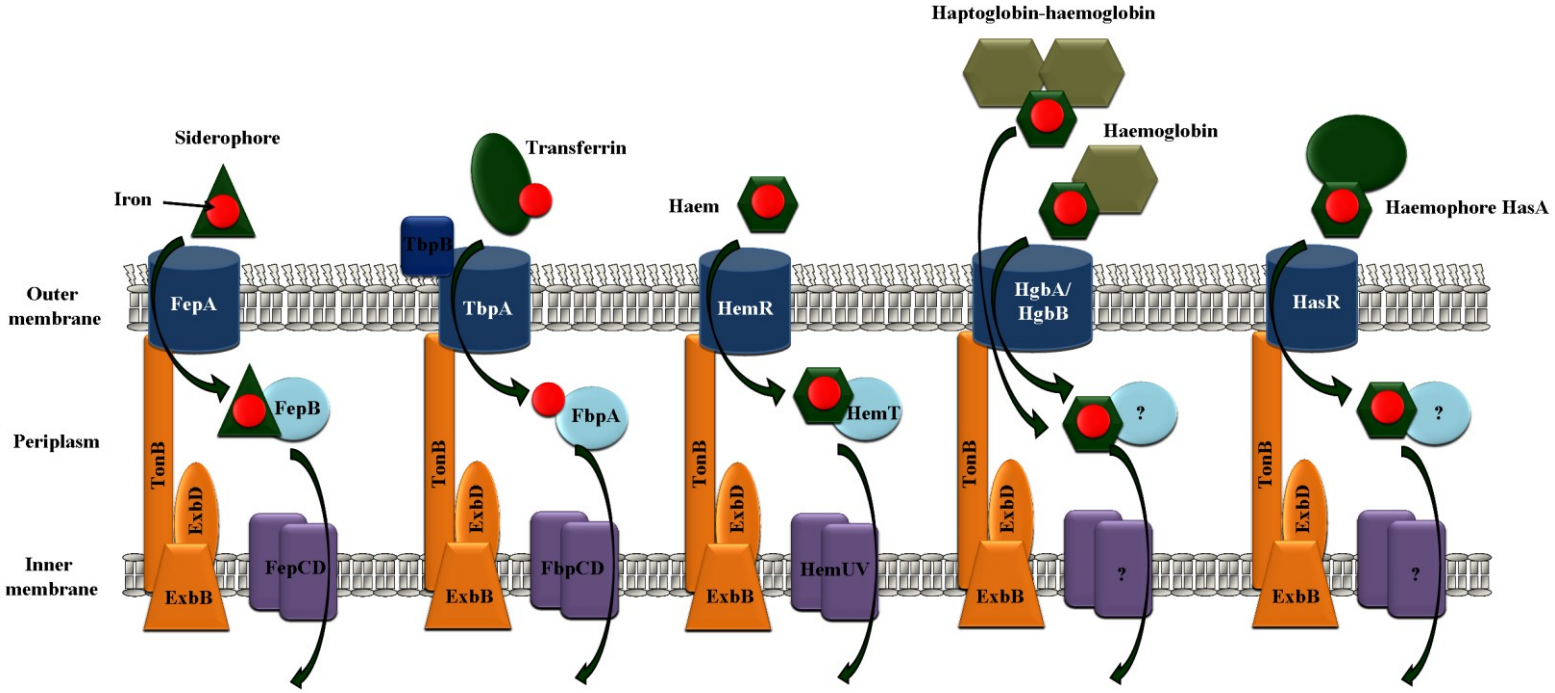


Figure 1-14 Different iron acquisition systems which involve TonB-dependent outer membrane receptors (siderophore receptor FepA in *E. coli*, transferrin receptor TbpAB in *Neisseria* species, haem receptor HemR in *Yersinia enterocolitica*, haemoglobin and haptoglobin-haemoglobin receptors HgbA and HgbB in *P. multocida* and haemophore receptor HasR in *Serratia marcensis*), periplasmic binding proteins (FepB, FbpA, HemT) and inner membrane ABC transporters (FepCD, FbpCD, HemUV) (Faraldo-Gómez & Sansom, 2003). The energy for these uptake mechanisms is transduced by the TonB-ExbB-ExBD complex.

With respect to the transmembrane β-barrel proteins, the orientation of proteins in membranes (OPM) database (http://opm.phar.umich.edu/superfamilies.php?class=2) categorized these β-barrel proteins into 25 superfamilies **(Table 1-2)**. This database is based on a computational approach for positioning proteins in the membranes (Lomize *et al.*, 2006). Classification was made at the level of superfamily based on evolutionary relatedness. Proteins with superimposable 3D structures and detectable sequence homology are further subdivided into families.

Table 1-2. Classification of transmembrane β-barrel proteins into superfamilies and families according to the OPM database (n = number of transmembrane strands) (http://opm.phar.umich.edu/superfamilies.php?class=2)

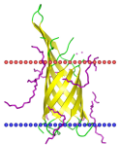
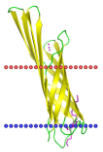
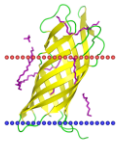
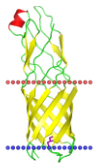
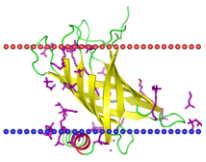
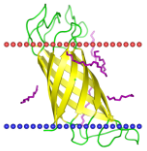
No	Superfamily	Example	Structure
1.	OmpA-OmpF porin family (n=8)		
	1.1 OmpA family	OmpA (*E. coli*)	
	1.2 Enterobacterial Ail/Lom protein	OmpX (*E. coli*)	
2.	Opacity protein (n=8)		
	2.1 Opacity porins	NspA (*N. meningitidis*)	
3.	OmpW family (n=8)		
	3.1 OmpW family	OmpW (*E. coli*)	
4.	Lipid A acylation (n=8)		
	4.1 Lipid A acylation	PagP (*E. coli*)	
5.	Lipid A 3-O-deacylase (n=8)		
	5.1 Lipid A 3-O-deacylase	PagL (*P. aeruginosa*)	

Table 1-2. (continued)

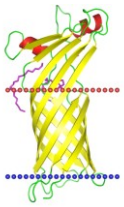
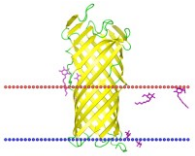
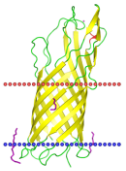
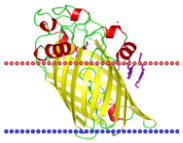
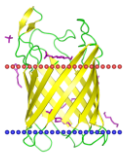
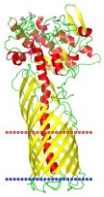
No	Superfamily	Example	Structure
6.	Major OMP (n=8)		
	6.1 Major OMP from Thermophilic eubacteria	Major OMP (*Thermus thermophilus*)	
7.	Omptin (n=10)		
	7.1 Outer membrane protease omptin	OmpT (*E. coli*)	
8.	Outer membrane adhesion protein (n=10)		
	8.1 Outer membrane adhesion protein OpcA	OpcA (*N. meningitidis*)	
9.	Lipid A deacylase (n=12)		
	9.1 Lipid A deacylase	LyxR (*S. enterica*)	
10.	Oligogalactoronate-specific porin (n=12)		
	10.1 Oligogalactoronate-specific porin (KdgM)	NanC (*E. coli*)	
11.	Autotransporter (n=12)		
	11.1 Autotransporters of N-terminal passenger domain	EstA (*P. aeruginosa*)	

Table 1-2. (continued)

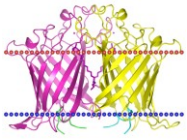
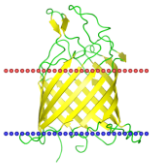
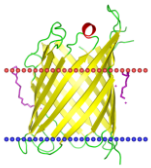
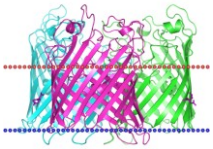
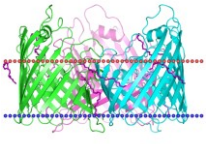
No	Superfamily	Example	Structure
12.	Outer membrane phospholipase (n=12)		
	12.1 Outer membrane phospholipase A	OmpLA (*E. coli*)	
13.	Nucleoside-specific channel-forming membrane porin (n=12)		
	13.1 Nucleoside transporter Tsx	Tsx (*E. coli*)	
14.	FadL outer membrane protein (n=14)		
	14.1 Fatty acid transporter FadL family	FadL (*E. coli*)	
15.	OmpG porin (n=14)		
	15.1 OmpG proin	OmpG (*E. coli*)	
16.	Trimeric porins (n=16)		
	16.1 General bacterial porin	OmpC (*E. coli*)	
	16.2 Rhodobacter PorCa porin	Porin (*Rhodobacter capsulatus*)	

Table 1-2. (continued)

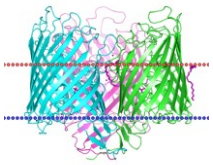
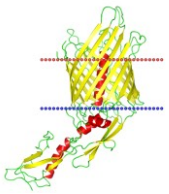
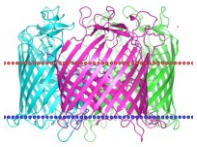
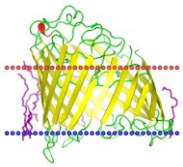
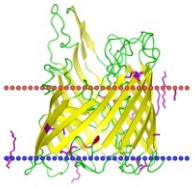
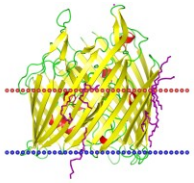
No	Superfamily	Example	Structure
	16.3 Porins O and P	OprP (*P. aeruginosa*)	
17.	Omp85-TpsB transporters (n=16)		
	17.1 Outer membrane protein insertion porin (OmpIP)	FhaC (*Bordetella pertussis*)	
18.	Sugar porins (n=18)		
	18.1 Maltoporin-like proteins	Maltoporin (*E. coli*)	
19.	OprD/AlgE superfamily (n=18)		
	19.1 Outer membrane porin Opr	OprD (*P. aeruginosa*)	
	19.2 Alginate export porin	Alginate export protein (*P. aeruginosa*)	
20.	Ligand-gated protein channels (n=22)		
	20.1 Brucella-Rhizobium porin	BtuB (*E. coli*)	

Table 1-2. (continued)

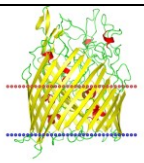
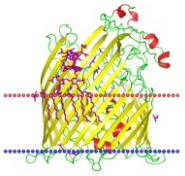
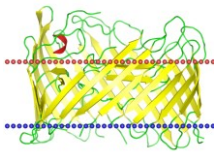
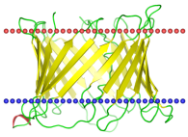
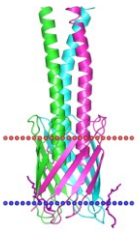
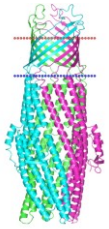
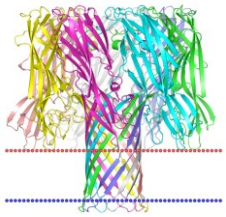
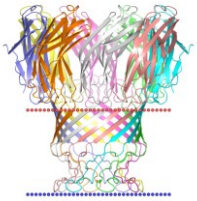
No	Superfamily	Example	Structure
	20.2 Outer membrane receptor (OMR)	FepA (*E. coli*)	
		FhuA (*E. coli*)	
		HasR (*Serratia marcescens*)	
21.	Fimbrial usher porin (n=24)		
	21.1 Fimbrial usher porin	FimD (*E. coli*)	
22.	Mitochondrial and plastid porins (n=19)		
	22.1 Voltage-dependent anion channel (VDAC)	VDAC-1 channel (*Homo sapiens*)	
23.	Trimeric autotransporter (n=12)		
	23.1 Autotransporter-2 (AT-2)	Hia (*H. influenzae*)	

Table 1-2. (continued)

No	Superfamily	Example	Structure
24.	Outer membrane factor (OMF) (n=12)		
	24.1 TolC-like proteins	TolC (*E. coli*)	
25.	Leukocidin-like		
	25.1 Alpha-hemolysin channel-forming toxin (n=14)	Alpha-hemolysin (*Staphylococcus aureus*)	
	25.2 Mycobacterial porin (MBP) (n=16)	MspA (*Mycobacterium smegmatis*)	

1.6.2.4 OMP profiles

The outer membrane proteome can be visualized by protein separation on either one-dimensional (1D) or two-dimensional (2D) sodium sodecyl sulphate (SDS) polyacrylamide gels. The principle of protein separation will be described in **Chapter 3**. The OMPs visualized on the SDS polyacrylamide gels, also called the OMP profiles, represent the outer membrane proteome under a particular growth condition and at a particular time point. The OMP profiles allow qualitative and quantitative comparisons of the outer membrane proteomes from different strains of bacteria and under different growth conditions. OMPs profiles have previously been used to study the outer membrane proteomes of many bacterial species including *Y. pestis* (Pieper *et al.*, 2009), *E. coli* (Alteri & Mobley, 2007), *Caulobacter crescentus* (Phadke *et al.*, 2001), *Edwardsiella tarda* (Kumar *et al.*, 2009), *Prevotella intermedia* (Yu *et al.*, 2007), *Campylobacter jejuni* (Hobb *et al.*, 2009) and *Legionella pneumophila* (Khemiri *et al.*, 2008). OMP profiles were previously used to examine outer membrane protein diversity and to classify strains of *P. multocida* recovered from different host species (avian, bovine, ovine and porcine origins) based on the molecular mass variation of the two major proteins, OmpA and OmpH, and by variation of the minor OMPs (Davies *et al.*, 2003a; 2003b; 2003c; 2004). Boyce *et al.* (2006) compared the OMP profiles of *P. multocida* grown under *in vitro* and *in vivo* growth conditions.

1.7 Objectives

The main aim of this PhD project was to compare and contrast the outer membrane proteomes of *P. multocida* isolates associated with diseased cattle, sheep, pigs and poultry using appropriate prediction and identification methods. The study will provide an improved understanding of the outer membrane proteomes of *P. multocida* isolates from different host species and the adaptation of this bacterium to different animal hosts.

First, bioinformatic prediction workflow was designed by using multiple predictor programs in association with consensus prediction and manual confirmation methods to confidently predict putative OMPs and estimate the size of the outer membrane proteome from the available genome sequences of *P. multocida*. These putative OMPs were functionally classified and studied in detail.

Second, comparative proteomic methods were used to experimentally validate and characterize the predicted outer membrane proteome of *P. multocida* isolates associated with diseased cattle, sheep, pigs and poultry. Different OMP extraction methods were tested and the optimal method was selected. The OMPs were separated using appropriate protein separation methodologies and a combination of proteomic identification methods were employed to determine the identities of the OMPs. These allowed a comparison of the outer membrane proteomes of *P. multocida* isolates associated with diseased cattle, sheep, pigs and poultry leading to an understanding of host adaptation process and host-pathogen interaction.

Third, comparative studies of the outer membrane proteomes of the same isolates under different growth conditions were performed. These included different stages of the growth, different rates of aeration, growth under iron-limited condition, growth in serum and in combinations of media and serum, and growth on plates and as biofilms. Growth under these various conditions were examined to maximize the number of identified proteins and to understand changes in the outer membrane proteome in response to different growth conditions. The response of different *P. multocida* isolates associated with different diseased animal hosts to these various *in vivo*-like growth conditions will improve our understanding of the interactions between this bacterium and its animal hosts.

Chapter 2: Prediction of the *P. multocida* outer membrane proteome

2.1 Introduction

The Gram-negative bacterium *Pasteurella multocida* is responsible for economically significant infections of a wide range of animal species. The organism causes a variety of disease syndromes which include pneumonic pasteurellosis of ruminants and pigs, porcine progressive atrophic rhinitis (PAR), fowl cholera, bovine haemorrhagic septicaemia, and human infections via carnivorous bites or scratches (Harper *et al.*, 2006). Like all Gram-negative bacteria, the cell envelope of *P. multocida* consists of a symmetrical inner membrane and an asymmetrical outer membrane, separated by the periplasmic space and peptidoglycan layer (St Michael *et al.*, 2005). The outer membrane consists of an inner phospholipid layer and an outer LPS leaflet (Cowan, 2012). It serves as a selective barrier that controls the passage of nutrients and waste products into and out of the cell and is the interface between pathogen and host. The outer membrane harbours two classes of proteins, integral membrane proteins and peripheral lipoproteins, which together account for 2-3% of the total encoded proteins (Lin *et al.*, 2002). Integral membrane proteins typically have a β-barrel structure whereas lipoproteins are mostly anchored to the inner leaflet of the outer membrane (Ruiz *et al.*, 2006). The biosynthesis and translocation of these two groups of proteins have previously been reviewed in **Chapter 1**. Outer membrane proteins (OMPs) play varied and important roles for bacteria, allowing them to adapt to different environments and host niches (Ruiz *et al.*, 2006). These roles include biogenesis and integrity of the outer membrane, nonspecific porin activity, energy-dependent transport, adherence and membrane-associated enzymatic activity (Lin *et al.*, 2002). In *P. multocida*, certain OMPs play important roles in virulence and have been utilized as vaccine antigens (Dabo *et al.*, 2007).

2.1.1 The genome and proteome of P. multocida

The genome is a store of biological information essential for the construction and maintenance of living organisms. The genomic information is transcribed into the transcriptome by enzymes and other proteins. The transcriptome comprises messenger RNA (mRNA) molecules encoded from their associated genes as required by the organism (Brown, 2007). The mRNA molecules are translated into proteins which are able to carry out biological functions. The translation of a collection of proteins at a particular time is named the proteome (Brown, 2007). The genome information is obtained by genome sequencing and the gene location can be computationally determined. The protein-coding genes consist of open reading frames (ORFs) which contain an initiation codon, a series of codons encoding for the proteins and a termination codon. Since the codon is a triplet code, genome scanning for the gene location has to be done in six reading frames, three forward and three reverse directions (Brown, 2007).

In 2001, May *et al.* (2001) sequenced the first complete genome (NC_002663) of a *P. multocida* strain (Pm70) by shortgun sequencing. Strain Pm70 is an avian capsular type A serotype 3 strain associated with fowl cholera. The genome of Pm70 consists of a single circular chromosome consisting of 2.3 megabase pairs and contains 2,092 genes, 2,015 predicted ORFs, six ribosomal RNA operons and 57 tRNAs. This genome encodes a number of enzymatic pathways including oxidative pentose phosphate and Entner-Doudoroff pathways, glycolysis, gluconeogenesis, trichloroacetic acid (TCA) cycle, sulphur uptake and metabolism pathways, and nitrogen and folic acid metabolism pathways (May *et al.*, 2001). The authors identified 53 ORFs involved in iron acquisition and two adhesin-encoding genes, named *Pasteurella* filamentous haemagglutinin (*pfh*)*B1* and *pfhB2*, which were homologous to the filamentous haemagglutinin gene *fhaB* in *B. pertussis* (May *et al.*, 2001).

2.1.2 Introduction to bioinformatic predictions and algorithms

Bioinformatics is a merging between conceptualizing biology in terms of molecules and applying information technologies derived from applied maths, computer science and statistics to understand and manage the bioinformation associated with these molecules on a large scale (Luscombe *et al.*, 2001). Bioinformatics provides tools to analyze biological sequence data, genome content and arrangement, and to predict structure and function of these molecules (Luscombe *et al.*, 2001). Bioinformatic prediction prior to experimental characterization of these molecules can reduce cost and time spent on the experiments. However, these data are complex and massive in size. Bioinformatic analyses often require computational methods for data storage such as the nucleic acid database GenBank and the universal protein resource UniProt, data organization and integration such as the National Centre for Biotechnology Information NCBI, and methods for understanding biological data such as sequence alignment, homology searches, searching for functional domains, predictions of structure, function and localization, and large-scale analyses of genomes, transcriptomes, metabolomes and proteomes (Luscombe *et al.*, 2001). These databases and programs have been created based on various problem-solving algorithms. The algorithm is a sequence of steps that are systematically executed to produce the desired outcome (Keedwell & Narayanan, 2005). In this introduction, different algorithms for the prediction and characterization of OMPs are described.

2.1.3 OMP predictors

The majority of OMPs can be bioinformatically differentiated and predicted by using their amino acid compositions (Gao *et al.*, 2010; Gromiha & Suwa, 2006a; Gromiha, 2005), specific protein modifications and sorting mechanisms (Fariselli *et al.*, 2003; Juncker *et al.*, 2003), and unique sequences and structural patterns (Emanuelsson *et al.*, 2007; Jackups *et al.*, 2006; Mirus & Schleiff, 2005; Valavanis *et al.*, 2006; Waldispuhl *et al.*, 2006). Predictors of outer membrane-located proteins employ a variety of algorithms and methods

having different accuracies and sensitivity levels of prediction (Bagos *et al.*, 2004b; Berven *et al.*, 2004, 2006; Bhasin *et al.*, 2005; Bulashevska & Eils, 2006; Fyshe *et al.*, 2008; Gardy *et al.*, 2005, 2003; Garrow *et al.*, 2005a; Gromiha & Yabuki, 2008; Gromiha *et al.*, 2007; Hu & Yan, 2008; Imai *et al.*, 2008; Ou *et al.*, 2008; Remmert *et al.*, 2009; Shen & Chou, 2010; Szafron *et al.*, 2004; Wu *et al.*, 2007; Yan *et al.*, 2008; Yu & Lin, 2004; Yu *et al.*, 2006, 2010; Zhai & Saier, 2002). These predictors can be categorized into three groups: (1) subcellular localization or global predictors which can differentiate between proteins from different compartments; (2) transmembrane β-barrel protein predictors which distinguish β-barrel structures from transmembrane α-helical proteins predominantly found in the inner membrane; and (3) lipoprotein predictors which can discriminate between inner membrane and outer membrane lipoprotein signal peptides (Díaz-Mejía *et al.*, 2009).

2.1.3.1 Prediction of subcellular localization

Subcellular localization is an important characteristic of proteins. This approach can categorize proteins into different compartments including cytoplasm, inner membrane, periplasmic space, outer membrane and extracellular. Examples of the predictors belonging to this group are described below.

Proteome Analyst, which predicts subcellular localization along with explanations (genome-wide and proteome-wide annotations), and **PA subcellular localization** (PA-SUB), which predicts only subcellular localization, were developed by the use of similarity searches on sequence data, extraction of text annotations from homologs followed by the use of text features as classifiers (Naïve Bayes or NB classifiers) for the machine learning algorithm (Lu *et al.*, 2004; Szafron *et al.*, 2004). Comparison to PSORTb (see below) suggested that PA predicted OMPs at a better sensitivity (94.7% in PA and 90.3 in PSORTb) and accuracy (92% in PA and 90.3% in PSORTb), but with comparable precision (98.6% in PA and 98.8% in PSORT-B) (Lu *et al.*, 2004; Szafron *et al.*, 2004).

PSORTb is a predictive program of the five subcellular localizations of Gram-negative bacteria. This program combines a variety of individual predictors (Gardy *et al.*, 2003). These include SCL-BLAST for homology analysis of proteins of known localization; motif-based analysis for localization-specific motifs; detection of alpha-helical transmembrane regions of inner membrane proteins by HMMTOP; identification of OMPs by their beta-barrel structure; and overall amino acid composition-based subcellular localization prediction by SubLoc, which was later replaced by a new Support Vector Machine (SVM)-based method and signal peptide predictor. The overall precision and accuracy of PSORTb is 97% and 75%, respectively (Gardy *et al.*, 2003). Gardy *et al.* compared PSORTb to PA and CELLO by using proteins not included in any of these programs' training data and revealed that PSORTb achieved the highest precision of 98% over PA (87.5%) and CELLO (71.5%), and comparable recall to PA (Gardy *et al.*, 2005). However, PSORTb did not discriminate between β-barrel proteins and lipoproteins.

CELLO is a single-module program developed by using a machine learning algorithm named SVM based on multiple *n*-peptide composition to predict five subcellular localizations including cytoplasm, inner membrane, periplasmic space, outer membrane and extracellular space of Gram-negative bacteria (Yu & Lin, 2004). The overall accuracy of this program is 89% and is 14% higher than that of PSORTb. CELLO v.2 was developed based on a two-level SVM system: the first level contained features derived from the sequence; the second level used decisions of possible subcellular localization (Yu *et al.*, 2006). The authors compared their method to the homology search method. They suggested that if the training data set contained proteins with high homology levels, these would lead to the biased assessment of the performance of the program.

SOSUI-GramN was developed based on the physicochemical parameters of the N- and C-terminal signal sequences, and the total amino acids without the requirement of the

homology data of the known sequences (Imai *et al.*, 2008). SOSUI-GramN consists of three layers of filters: the first layer for discriminating the inner membrane proteins from others; the second layer for the classification of the non-inner membrane proteins into proteins involved in the Sec-dependent or Sec-independent pathways; and the third layer for further classification into cytoplasmic, periplasmic, outer membrane and extracellular proteins. This program predicts all five subcellular localizations of Gram-negative bacteria with precision and recall of 92.9% and 86.7%, respectively. This program also shows improvement in precision and recall of the extracellular proteins compared to PSORTb and CELLO.

SubLoc is a prediction program developed for subcellular localization based on SVM which is an effective method for supervised pattern recognition (Hua & Sun, 2001). The SubLoc program predicts the localization of proteins by recognition of their amino acid compositions. For prokaryotes, the overall accuracy of prediction is 91.4% for three subcellular localizations (cytoplasmic, periplasmic and extracellular).

SignalP is an amino acid sequence-based program that predicts secretory signal peptides and can be used for determining subcellular localization (Emanuelsson *et al.*, 2007).

2.1.3.2. Prediction of transmembrane β-barrel proteins

As the transmembrane β-barrel structure is one of the key characteristics of integral OMPs, various predictors have been developed based on this feature (Wimley, 2002). The general β-barrel OMPs consist of an even number of β-strands (from 8 to 22 strands) comprising monomeric, dimeric or trimeric barrels. Polar residues are located internally, while hydrophobic residues are exposed to the lipid-interacting outer surface. Diad repeats of hydrophobic and hydrophilic residues of the β-strands have been observed. There is a band of aromatic residues close to the bilayer interface (Wimley, 2002).

BOMP, β-barrel outer membrane protein predictor, uses two separate components: recognition of a C-terminal pattern typical of many transmembrane β-barrel proteins and calculation of a β-barrel score of the sequence based on amino acid patterns of known transmembrane β-strands (Berven *et al.*, 2004). This program classifies the predicted proteins into five groups (one to five): group one for the highest confidence and group five for the lowest. The precision and recall of prediction by BOMP is 80% and 88%, respectively. However, BOMP can not predict OMPs which have less than eight transmembrane β-strands.

TMB-Hunt utilizes a modified *k*-nearest neighbour (*k*-NN) algorithm which is a simple instance-based learning algorithm based on the whole sequence amino acid composition and evolutionary information to discriminate between transmembrane or non-transmembrane β-barrel proteins (Garrow *et al.*, 2005a). This program achieves an accuracy of 92.5% and sensitivity of 91%.

TMBETADISC-RBF was developed by using radial basis function networks (RBF) and position specific scoring matrix (PSSM) profiles generated by PSI-BLAST and non-redundant protein database (Ou *et al.*, 2008). This program can discriminate the transmembrane β-barrel OMPs from other types of proteins with the highest accuracy of 96.4%.

MCMBB is a program that discriminates transmembrane β-barrel proteins from globular and transmembrane α-helical proteins based on the first order Markov Chain model which identifies the alternating pattern of hydrophobic-hydrophilic residues in the transmembrane spanning β-strands (Bagos *et al.*, 2004a).

PredTMBB is performed using Veterbi, N-best and posterior decoding with dynamic programming algorithms to identify transmembrane strands and loop regions (Bagos *et al.*,

2004b, c). This program correctly predicted 9 to 10 of 14 known-structure OMPs with per residue accuracy of 84-88% and the program provided a score used to identify β-barrel OMPs below the threshold of 2.995.

ProfTMB is a profile-based Hidden Markov Model (HMM) that uses a Z-value to indicate transmembrane β-barrel structures and provides four-state secondary structure prediction including up-strand, down-strand, periplasmic hairpin and extracellular loop, with a whole-protein discrimination accuracy of 70% and per-residue accuracy of 86% (Bigelow & Rost, 2006).

Zou and Wang improved the HMM-based β-barrel transmembrane protein topology predictor using either two-state (strand and non-strand) or four-state (up-strand, down-strand, inner loop and outer loop) predictions (Zou & Wang, 2007). Using 26 non-redundant known-structure OMPs, their method outperformed PROFtmb and PredTMBB at accuracies of 88% and 89.7% for two-state and four-state predictions, respectively.

Bagos *et al.* (2005) evaluated the performance of different methods including HMMs (e.g. HMM-B2TMR, PROFtmb, PredTMBB and BETA-TM), Neural Networks (NNs) and SVMs (e.g. TMBETA-NET, B2TMPRED, PSIPRED and TBBPRE), for transmembrane β-barrel protein prediction using 20 known transmembrane β-barrel OMPs as a training set and concluded that the HMM-based methods (HMM-B2TMR, ProfTMB and PredTMBB) were the best predictors. PredTMBB had the highest score and was used to evaluate the performance of secondary structure prediction methods. This program performed better than HMM-B2TMR and ProfTMB, respectively. The HMM-based programs could avoid the prediction of signal sequences as transmembrane strands. The authors also suggested that consensus prediction methods performed better than a single program.

TMBETA-NET was developed to identify transmembrane β-barrel OMPs based on the analysis of amino acid compositions of OMPs using statistical methods and NN algorithms (Gromiha & Suwa, 2005; Gromiha *et al.*, 2005). The program could correctly predict OMPs with an average accuracy of 80-95%. Interestingly, the authors observed that the composition of Ser, Asn and Gln were higher in OMPs and are believed to play a role in stability and the formation of β-barrel structures. On the other hand, Glu, His, Ile and Cys were enriched in globular proteins. Gromiha and Suwa (2006b) compared the performance of different machine learning algorithms, e.g. Bayes functions, NNs, logistic functions, SVMs, regression analysis, nearest neighbour methods, metalearning, decision trees and rules for differentiating OMPs and non-OMPs based on amino acid composition and available sequences. They suggested that the performance of NN-based methods on 1,088 protein data sets (208 OMPs and 880 non-OMPs) achieved higher accuracy (91%) and comparable specificity and sensitivity to others.

Söding (2005) developed a predictive program for protein homology prediction and classification of OMPs based on the use of HMM-HMM comparison and the β-barrel structure prediction method. Taylor *et al.* (2006) developed a prediction method based on Bayesian Networks to identify β-barrel transmembrane proteins. Their method accurately predicted the β-barrel strands from non-strands with an accuracy of 88.6%, whereas the prediction of overall protein topology achieved only 42.7% accuracy.

PROB is another programme developed for the prediction of β-barrel membrane proteins and 114 proteins of *Mycobacterium tuberculosis* were predicted (Pajón *et al.*, 2006). The authors compared their predictor with TMB-Hunt and BOMP and revealed a few proteins predicted by the three.

TMBpro was developed for the prediction of secondary structure and tertiary structure of β-barrel transmembrane proteins (Randall *et al.*, 2008). The performance of the program was higher than PredTMBB and transFold by at least 4 %.

Although many programs have been developed based on various methods, the use of multiple programs may provide better coverage and higher confidence than the use of individual predictors. The numbers of well-annotated or known-structure OMPs are crucial for an improvement in the accuracy of predictors. As there are still a large number of Gram-negative bacterial proteins that are annotated as probable or putative OMPs, these transmembrane β-barrel OMP predictors will help in the identification of these proteins in genome derived proteomes. The user should also carefully interpret the predicted results to avoid including false positive and excluding false negative results.

2.1.3.3. Prediction of lipoproteins

To bioinformatically characterize the complete outer membrane sub-proteome, lipoproteins are another group of proteins to be considered. Lipoproteins contain signal peptides which consist of three regions: n-region, h-region and c-region, previously described in **Chapter 1**. The process of lipid modification was also explained in **Chapter 1**. Lipoproteins are found in both the inner and outer membranes of Gram-negative bacteria. Lipoproteins located in the cytoplasmic membrane have an aspartate residue at position +2 (D+2) after the cysteine, while lipoproteins located in the outer membrane have other residues in this position except phenylalanine, tryptophan, tyrosine, glycine and proline (Seydel *et al.*, 1999). The signal peptides (15-30 amino acids) of the modified prolipoproteins are cleaved by signal peptidase II (SPaseII) with a cleavage site of (LVI)(ASTG)(GA)↓C (lipobox), also called lipoprotein signal peptidase (Lsp) (Juncker *et al.*, 2003). Signal peptidase I (SPaseI) processes signal peptides of secreted extracellular proteins.

LipoP was developed based on HMM and NN to discriminate between lipoproteins, SPaseI-cleaved proteins, cytoplasmic and transmembrane proteins with an accuracy of 96.8% (Juncker *et al.*, 2003). This program also separates the N-terminal transmembrane helices from the signal peptides. The authors compared their predicted lipoproteins to the experimental data obtained from 12 Gram-negative strains and found that they could correctly identify 94.6% of the predicted lipoproteins.

Lipo is a program that discriminates outer membrane lipoproteins from non-lipoproteins in Gram-negative bacteria based on a regular pattern specific to the lipobox recognized by signal peptidase II (Berven *et al.*, 2006). This program groups the predicted lipoproteins into three groups: low, medium and high confidence, based on the precision score. The highest precision of this program was 92%. The authors suggested that this program could be an alternative to LipoP because it was based on different prediction methods. Both programs can be used in combination to increase the reliability of the prediction as well as to obtain a better coverage of the lipoproteins. The authors confirmed this by predicting outer membrane lipoproteins of *Methylococcus capsulatus* using both LipoP and Lipo. They found 91 proteins were predicted by Lipo and 63 proteins were predicted by LipoP, whereas 56 lipoproteins were predicted by both programs (Berven *et al.*, 2006).

2.1.3.4 Evaluation of predictor performance

Newly developed bioinformatic predictors have to be evaluated for their performance before making them available to users. The evaluation process can be achieved in a number of methods such as cross-validation and jack-knife (Rubingh *et al.*, 2006). Cross-validation is a statistical evaluation model which divides data into two groups: one used to train a model and the other used to test the model (Rubingh *et al.*, 2006). For example, in the five-fold cross-validation, the data will be divided into five groups within each iteration: four groups used for training and one used for testing (Gardy *et al.*, 2005; Ou *et*

al., 2008). Jackknifing is another evaluation method in which one group of data is left out each time and the remaining groups are used for training the model, until each group has been left out once (Rubingh *et al.*, 2006). These methods will provide statistical parameters such as accuracy, recall/sensitivity, precision and specificity which are used to validate and compare different models. Accuracy is a proportion of true results in the population. Precision is a proportion of the true positives among all positive results. Recall/sensitivity is a proportion of the true positives which are correctly identified. On the other hand, specificity is a proportion of the true negatives which are correctly identified.

2.1.3.5 Combination of different predictors

Bioinformatic predictors have been used to identify OMPs in several Gram-negative bacterial species (Berven *et al.*, 2006; Boyce *et al.*, 2006; Chung *et al.*, 2007; Huntley *et al.*, 2007). Chung *et al.* (2007) used five predictors (PA, PSORTb, BOMP, Lipo and LipoP) to scan the *A. pleuropneumoniae* genome and predicted 93 OMPs of which 50% were identified by proteomic analysis. Babujee *et al.* (2007) estimated the outer membrane proteome of *Dickeya dadantii* with the use of PSORTb, BOMP, Lipo, LipoP, SignalP and PredTMBB. However, disagreement between the predicted results from individual programs is frequently observed. A combination of different predictors, together with consensus prediction, has been shown to increase the coverage and accuracy of the predicted outer membrane proteome (Díaz-Mejía *et al.*, 2009; Viratyosin *et al.*, 2008) including that of transmembrane β-barrel proteins (Bagos *et al.*, 2005). Heinz *et al*. also employed a manual confirmation step to remove false positives and increase the confidence of the predicted outer membrane proteome (Heinz *et al.*, 2009).

In a previous *P. multocida* study (Al-Hasani *et al.*, 2007), three predictors (two subcellular localization and one lipoprotein) were used to predict 129 proteins as secreted, outer

membrane or lipoprotein from the publicly available genome of *P. multocida* avian strain Pm70 (May *et al.*, 2001). However, certain predicted proteins were not confirmed as OMPs by all three predictors. Boyce *et al.* (2006) identified 35 proteins by proteomics from the *P. multocida* avian strain X-73 but only one third of these proteins were predicted to be OMPs by a combination of two subcellular localization predictors. Therefore, our understanding of the outer membrane proteome of *P. multocida* remains elusive.

2.1.4 Prediction of physicochemical properties of OMPs

Physicochemical properties include basic protein information that can be predicted from a protein sequence and include amino acid composition, molecular mass, pI, secondary structures and hydrophobicity. This section will explain certain physicochemical properties of the OMPs.

2.1.4.1 Predictions of numbers of transmembrane β-barrel and α-helical segments

Transmembrane strands, both β-strands and α-helices, are basic structural units used to discriminate transmembrane β-barrel and transmembrane α-helical proteins by various predictors. The properties of the transmembrane α-helices were described by Arkin and Brunger (Arkin & Brunger, 1998). The transmembrane α-helix is a stretch of 15-30 hydrophobic amino acids (e.g. alanine, valine, leucine, isoleucine, proline, methionine, phenylalanine and tryptophan) which form a helical structure by hydrogen bonding in the phospholipid bilayer. Aliphatic amino acids are distributed randomly within the helices whereas aromatic amino acids are commonly found at the ends of the helices. The average hydrophobicity of the transmembrane α-helices is very high. The transmembrane α-helices can be divided into two groups based on the number of transmembrane α-helices. Bitopic transmembrane α-helical proteins contain a single transmembrane α-helix, whereas polytopic transmembrane α-helical proteins contain multiple α-helices. More polar amino acids are found in polytopic proteins compared to bitopic proteins because the

transmembrane α-helices in polytopic proteins interact with other helices and the lipid environment. The interaction between the transmembrane α-helices can accommodate polar amino acids within the lipid environment.

The transmembrane β-strands usually form antiparallel sheets of β-barrel structure in the outer membrane. The properties of the transmembrane β-strands were reviewed by Garrow *et al.* (2005b). The β-strands consist of an inside-outside dyad repeat motif of alternating hydrophobic residues, facing the lipid layer, and polar or intermediate polar residues, facing the inside of the β- barrel. Each strand has an average of 6-22 amino acids, with 12 residues most frequently found. Similar to the transmembrane α-helices, aromatic amino acids are found at both ends of the β-strands in order to maintain the protein stability within the membrane.

2.1.4.2 Hydrophobicity prediction

Hydrophobicity of the proteins can be predicted by an average hydrophobicity value of all amino acids within the protein sequence (Kyte & Doolittle, 1982). This is an important property for the identification of membrane-spanning regions and the prediction of membrane proteins. **Figure 2.1** shows hydrophobicity scores of all 20 amino acids based on the scale developed by Kyte and Doolittle. However, different hydrophobicity scales have been created for various purposes such as the Engelman scales (Engelman *et al.*, 1986) and Eisenberg (Eisenberg *et al.*, 1984) for predicting transmembrane regions, and the Rose scales (Rose *et al.*, 1985) and Janin (Janin, 1979) for predicting globular proteins.

2.1.4.3 Prediction of molecular mass and p*I*

Molecular mass and isoelectric p*H* (p*I*) are basic properties of all proteins. The molecular mass of any protein is the sum of the average isotopic masses of the amino acids in that protein. The p*I* value is calculated from an average of the p*K* values of the amino acids in the protein (Gasteiger *et al.*, 2005).

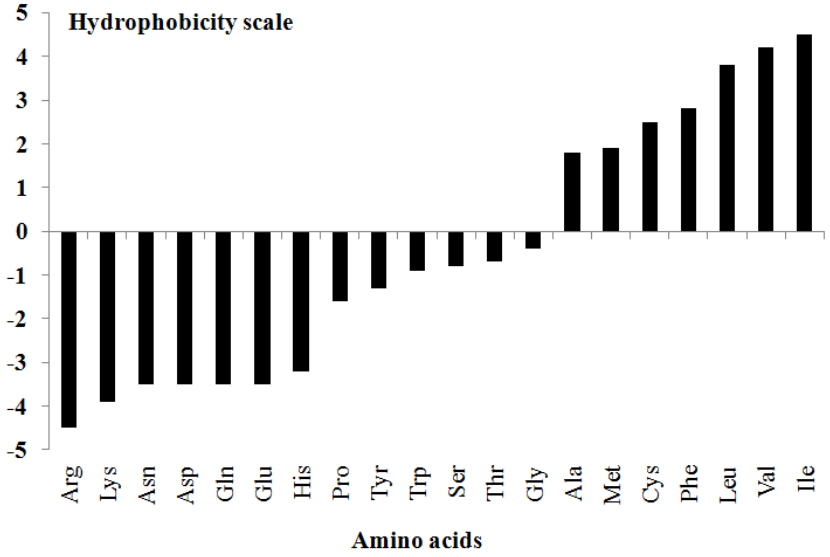


Figure 2-1. The hydrophobicity score of 20 amino acids. The x-axis shows amino acids and the y-axis shows the hydrophobicity scale.

2.1.5 OMP databases

As information grows on the structure and function of membrane proteins, including OMPs, various research groups have created public databases which systematically store protein information and provide useful resources for these proteins. This section provides examples of these OMP and OMP-related databases (**Table 2-1**).

The **OPM** (Orientations of Proteins in Membranes) database stores a collection of transmembrane and peripheral protein data and calculates the arrangements of these proteins within the membrane compared to experimental data (Lomize *et al.*, 2006). This database classifies membrane proteins into transmembrane proteins (α-helical bitopic and polytopic, and β-barrel proteins), peripheral or integral monotopic proteins (all α-, all β-, α/β and α+β) and peptides. Within this database, 86 β-barrel structures have been deposited and classified into 25 superfamilies as shown in **Table 1-2**. Most of them belong to Gram-negative bacteria, but a few were obtained from *Mycobacterium*, *Staphylococcus aureus*, mitochondria and plastids.

OMPdb is a database for transmembrane β-barrel OMPs from Gram-negative bacteria (Tsirigos *et al.*, 2011). This database stores 69,354 OMPs from 2,712 Gram-negative bacterial species which are classified into 85 families based on structure and function. The database allows cross-referencing to other databases. The developers of this database have claimed that OMPdb provides a complete classification and manually corrected annotation of the transmembrane β-barrel OMPs with detailed literature references.

The **TMFunction** database collects functional residues in membrane proteins reported in the literatures which are important for understanding the relationship between sequence, structure and function (Gromiha *et al.*, 2009).

Table 2-1. Outer membrane protein databases.

Database name	**Link**
OPM	http://opm.phar.umich.edu/
OMPdb	http://aias.biol.uoa.gr/OMPdb/
TMFunction	http://tmbeta-genome.cbrc.jp/TMFunction/
TCDB	http://www.tcdb.org/
DOLOP	http://www.mrc-lmb.cam.ac.uk/genomes/dolop/

TCDB (transport classification database) is a classification system based on both functional and phylogenetic information for membrane transport proteins, namely the transport classification (TC) system (Saier *et al.*, 2006, 2009). This TC classification system consists of five components including (1) channels and pores, (2) electrochemical potential-driven transporters, (3) primary active transporters, (4) group translocators and (5) transmembrane electron carriers. The transmembrane β-barrel OMPs are grouped in the channels and pores component which are further divided into 58 families.

DOLOP is a database which contains 278 different bacterial lipoproteins predicted from 234 bacterial genomes (Babu & Sankaran, 2002; Babu *et al.*, 2006). This database classifies lipoproteins into nine groups based on their functions: (1) structural proteins such as Omp P6, Omp16, Slp, NlpD, Lipoprotein E, Pcp and LolB (2) adhesins such as invasin InvH, (3) antigens such as Cag pathogenicity island protein, (4) enzymes such as MltA, MltB, MltC, MltD and GlpQ, (5) transporters such as a capsular polysaccharide exporter BexD, (6) binding proteins such as MetQ, (7) toxins such as entericidin A, (8) interesting but non-classifiable factors such as VacJ and RlpB and (9) hypothetical lipoproteins. The DOLOP database predicts 53 lipoproteins in *P. multocida*, accounting for 2.63% of the proteome (Babu & Sankaran, 2002).

Additionally, other membrane protein databases that are useful for the study of OMPs include the protein data bank RCSB (Berman *et al.*, 2000), the protein families database Pfam (Finn *et al.*, 2010), the topology database of transmembrane proteins TOPDB (Tusnády *et al.*, 2007), the database of protein subcellular localization PSORTdb (Yu *et al.*, 2011), the porin database server PRNDS (Katta *et al.*, 2004), and the protein data bank of transmembrane proteins PDBTM (Tusnady *et al.*, 2005).

2.1.6 Prediction of protein structures and functions

After the prediction of subcellular localization, detailed bioinformatic analyses of the predicted proteins help in the understanding of protein functions and of the relationship between sequence and structure. This section gives examples of such bioinformatic analyses including homology searches, prediction of protein structure and function, and prediction of protein-protein interaction.

2.1.6.1 Homology searches

Homology search is the identification of relatedness between DNA or protein sequences. The sequences are homologous if they share a common evolutionary ancestor (Pevsner, 2009). Homologous sequences from different species which have evolved from a common ancestor are called orthologs. If the sequences have evolved by gene duplication within the same species, these are called paralogs. The level of relatedness of the sequences can be quantitatively assessed by using identity and similarity. Similarity is used when the compared sequences do not have identical residues but they share similar biochemical properties (Dear, 2007). Homology searches can be performed by two types of sequence alignments: pairwise alignment for two sequences and multiple alignment for more than two sequences, and executed based on two alignment algorithms; global and local alignments (Pevsner, 2009). The global alignment uses the entire sequence, whereas the local alignment focuses on regions with the highest similarity. Most homology search programs use local alignment algorithms such as BLAST (basic local alignment search tool).

The BLAST program allows users to query a sequence to DNA or protein databases (Altschul *et al.*, 1990). The BLAST search can be used to determine orthologs and paralogs, to identify whether a gene or protein is present in a particular organism, to find

the identity of a sequence, to identify new genes, to investigate sequence variation and to examine residues that are important in the structure and function of the protein.

2.1.6.2 Prediction of protein structures

Protein structure can be described in four levels: primary structure for the amino acid sequence; secondary structure which is the formation of α-helices and β-sheets; tertiary structure which is the assembly and interactions of the α-helices and β-sheets; and quaternary structure which is the interaction of multiple subunits. The protein structural information is mainly deposited in the RCSB protein data bank PDB (Berman *et al.*, 2000) and the structural classification of proteins in the SCOP database (Murzin *et al.*, 1995). Sometimes, a protein with a single polypeptide chain can have multiple functional and structural units which are located on different parts of the sequence, called domains. A number of programs have been developed for the search of conserved domains using pattern-matching methods such as Pfam (Finn *et al.*, 2010), PROSITE (Sigrist *et al.*, 2010), SMART (Letunic *et al.*, 2009).

The structure of a target protein can be predicted by homology or comparative modelling which predicts the structure of the target protein by comparing with other related proteins with known sequence and structure (Pevsner, 2009). The quality of the predicted model depends on levels of sequence identity and similarity. Various programs are available for homology modelling such as 3D-JIGSAW (Bates *et al.*, 2001), Geno3D (Combet *et al.*, 2002), MODELLER (Eswar *et al.*, 2006), PredictProtein (Rost *et al.*, 2004) and SWISS-MODEL (Kiefer *et al.*, 2009). Other structural prediction methods include fold recognition (threading) and *ab initio* prediction. Fold recognition is used when lacking sequence matches between the model and the template, and the sequences are distantly related. The *ab initio* method predicts protein structure without using any models.

2.1.6.3 Prediction of protein functions

Protein functions can be predicted by a combination of various approaches (Dear, 2007). Proteins performing the same function are likely to have similar physicochemical properties. The protein function can be retrieved by aligning an unknown protein with its related proteins for which the function and structure are known. Another approach is the determination of functional domains which can reveal putative functions of the protein. Moreover, functional prediction can also be performed using the BLAST search. If the top matched proteins have a very high percentage of identity and good expectation (*E*) values, these top matches can be realigned again with the target protein (Altschul *et al.*, 1990). Once the common regions or domains are matched between these proteins, the annotation from the known proteins can be applied to the unknown one.

2.1.6.4 Prediction of protein-protein interaction

Many biological processes are accomplished by the interaction and association of proteins into stable or transient complexes forming a biological pathway. Predictions of protein-protein interactions are able to identify proteins that are functionally related and to assign putative functions to the uncharacterized proteins. Skrabanek *et al.* (2008) reviewed different approaches that have been used to predict protein-protein interactions: structural, genomic and biological approaches. The structural approach predicts protein-protein interaction based on existing protein structures allowing the determination of protein interacting sites and understanding the mechanisms of protein interaction. The genomic approach includes the prediction of co-localization or operon, and the occurrence of pairs of genes across multiple genomes. The biological approach is another method that predicts protein-protein interactions from available biological data such as gene expression analysis and microarray experiment. Examples of protein-protein interaction programs are STRING (Rost, 2001) and COGs (Tatusov *et al.*, 2000).

2.1.7 Prediction of protein expression levels

Protein expression levels can be predicted based on codon usage patterns which are varied in different organisms and different genes in the same organism (Wang *et al.*, 2001). Reasons for the varied codon usage patterns are due to overall nucleotide composition of the genome such as the GC-content; selective forces on highly expressed genes to enhance translational efficiency and horizontal gene transfer which transfers gene from different organisms (Supek & Vlahovicek, 2005). Different methods have been developed to predict protein expression levels based on the codon usage patterns including the Codon Adaptation Index (CAI) (Sharp & Li, 1987), F_{OP} which is a frequency of optimal codons within a gene (Ikemura, 1985), and the expression measure of a gene E(g) (Henry & Sharp, 2007).

2.1.7.1 Codon adaptation index (CAI)

The CAI measures the synonymous codon usage bias of the DNA and RNA sequences (Sharp & Li, 1987). It compares the synonymous codon usage of a target gene to the synonymous codon usage of a reference gene set. The CAI computes a weight for each codon from the reference sets and uses these weights to calculate the CAI value of each gene in the queried genome (Carbone *et al.*, 2005). This index ranges between zero and one. A value of one indicates that the best codons are being used for all amino acids in the protein. Certain proteins are known to be highly expressed such as ribosomal proteins and elongation factors (Carbone *et al.*, 2005; Mondal *et al.*, 2008).

2.1.8 Objectives

In this chapter, we used 10 different predictors classified into three groups (subcellular localization, transmembrane β-barrel protein and lipoprotein predictors) to identify putative OMPs from two available *P. multocida* genomes: the avian strain Pm70 and the unannotated genome of porcine non-toxigenic strain 3480. The predicted proteins in each

group were filtered by optimized criteria for the consensus prediction and the consensus predicted proteins from each group were integrated into a single list of proteins. We further incorporated a manual confirmation step which included a public database search against PubMed together with various sequence analyses, e.g. sequence and structural homology, conserved motifs/domains, functional prediction, and protein-protein interaction to enhance the confidence of prediction. Using these approaches, we were able to confidently predict the outer membrane proteomes of the two *P. multocida* strains.

2.2 Materials and methods

2.2.1 Gene prediction and annotation of the porcine strain genome of P. multocida

The publicly available genome of the avian *P. multocida* subsp. *multocida* strain Pm70 [GenBank: AE004439] and the unannotated genome of the porcine non-toxigenic *P. multocida* strain 3480 [Project ID: 32177] were used for all bioinformatic analyses. The avian strain genome containing 2,015 protein-coding genes was retrieved from NCBI. The 2,260 protein-coding genes of the unannotated porcine genome (kindly provided by Dr. A. Gillaspy, University of Oklahoma) were manually predicted using GeneMark (Besemer & Borodovsky, 2005) and automatically named using Blast2GO (Conesa *et al.*, 2005).

2.2.2 Selection of bioinformatic predictors

The scheme for the bioinformatic prediction of the OMPs is shown in **Figure 2-2**. Three groups of predictors, involving 10 genome-scale programs **(Table 2-2)**, were used to predict putative OMPs from the two genomes. Subcellular localization or global predictors included the programmes Proteome Analyst (Szafron *et al.*, 2004), PSORTb (Gardy *et al.*, 2005), CELLO (Yu & Lin, 2004) and SOSUI-GramN (Imai *et al.*, 2008); transmembrane β-barrel protein predictors included TMB-Hunt (Garrow *et al.*, 2005a), TMBETADISC-RBF (Ou *et al.*, 2008), BOMP (Berven *et al.*, 2004) and MCMBB (Bagos *et al.*, 2004b);

and outer membrane lipoprotein predictors included LIPO (Berven *et al.*, 2006) and LIPOP (Juncker *et al.*, 2003). Predicted results of each protein in the HTML or Excel formats from individual programmes were parsed using in-house built perl scripts **(Figure 2-3)**.

2.2.3 Analysis of agreement between pairs of bioinformatic predictors

The advantages of using multiple predictors over a single predictor can be evaluated by analysis of agreement between pairs of selected programs (Díaz-Mejía *et al.*, 2009). This analysis determines the level of agreement between different predictors by use of the following formula:

$$Agreement\ (A) \;=\; \frac{(P_1 \cap P_2)_L}{P'_L} \qquad \text{(Díaz-Mejía } et\ al.\text{, 2009)}$$

where $(P_1 \cap P_2)_L$ is the number of predicted proteins shared between predictor P_1 and P_2 for a subcellular location, L, and P'_L is the total number of predicted proteins from a lower coverage program of the predictor pair for that location. An agreement score (A) of one means that all proteins predicted by both methods (P_1 and P_2) are localized on a location, L. A score of zero means that there are no shared predicted proteins between the two predictors for a location, L. In-house built perl scripts were used to analyze the predicted results of each program before pairwise comparison and calculation of the agreement score **(Figure 2-4)**.

2.2.4 Criteria optimization

Predicted proteins from different programs in each group were filtered by consensus prediction with optimized criteria to eliminate redundancy and proteins of low confidence. These analyses were performed using Excel. We varied the criteria by increasing the number of positive predicted results as a threshold in each predictor group. For the subcellular localization and transmembrane β-barrel predictor groups, the criteria were varied from positive predicted proteins obtained by at least one, two, three or four

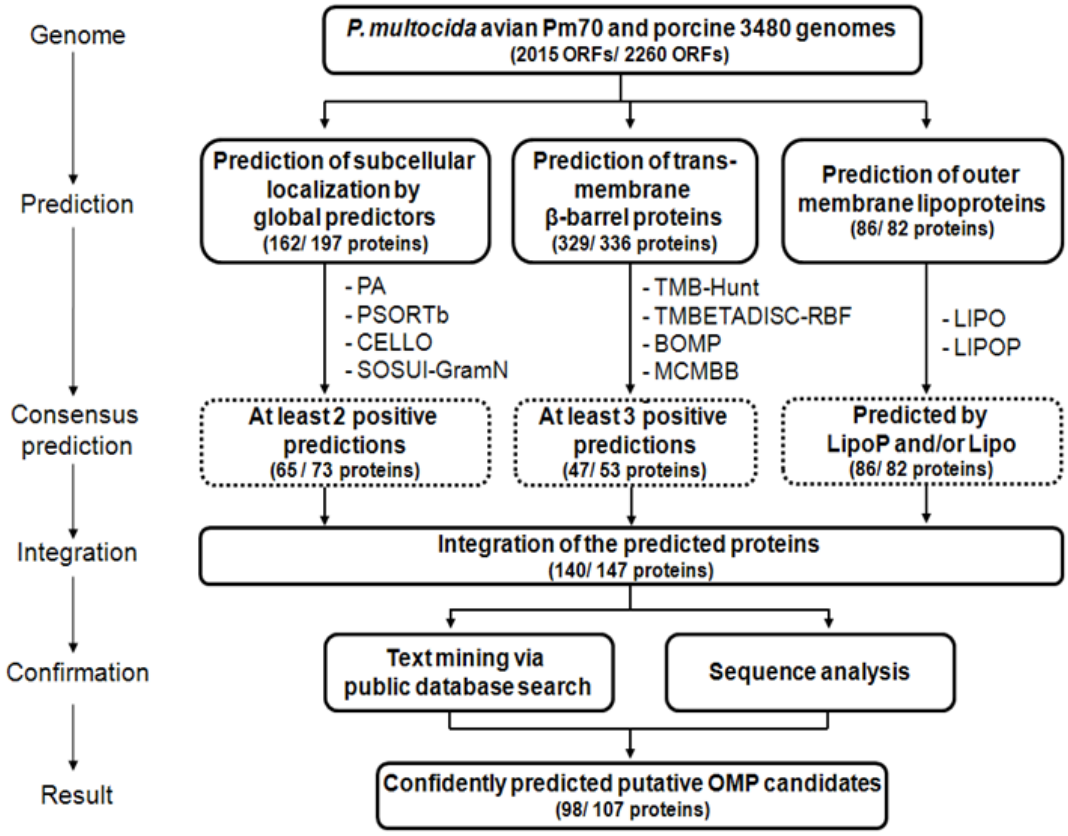


Figure 2-2. Diagram representing the workflow of bioinformatic prediction of putative OMPs from the avian and porcine strain genomes of *P. multocida*. Ten predictors were categorized into 3 groups: subcellular localization, transmembrane β-barrel protein prediction and outer membrane lipoprotein prediction. The predicted proteins in each group were filtered by consensus prediction and combined to form a single integrated list. Text mining and sequence analyses were used to confirm that the predicted proteins were outer membrane-associated with a high degree of confidence. The numbers of predicted proteins in each step are shown in parentheses: the first number represents proteins from the avian strain genome and the second number from the porcine strain genome.

Table 2-2. Bioinformatic predictors used for the OMP prediction.

Predictor group	Programme	Method of predictor	Reference
Subcellular localization	Proteome Analyst v. 3.0 (PA)	http://webdocs.cs.ualberta.ca/~bioinfo/PA/	(Szafron *et al.*, 2004)
	PSORTb v. 2.0	http://www.psort.org/psortb/	(Gardy *et al.*, 2005)
	CELLO v. 2.5	http://cello.life.nctu.edu.tw/	(Yu & Lin, 2004)
	SOSUI-GramN	http://bp.nuap.nagoya-u.ac.jp/sosui/sosuigramn/sosuigramn_submit.html	(Imai *et al.*, 2008)
Trans-membrane β-barrel structure	TMB-Hunt	http://bmbpcu36.leeds.ac.uk/~andy/betaBarrel/AACompPred/aaTMB_Hunt.cgi	(Garrow *et al.*, 2005a)
	TMBETADISC-RBF	http://rbf.bioinfo.tw/~sachen/OMPpredict/TMBETADISC-RBF.php	(Ou *et al.*, 2008)
	BOMP	http://services.cbu.uib.no/tools/bomp	(Berven *et al.*, 2004)
	MCMBB	http://athina.biol.uoa.gr/bioinformatics/mcmbb/	(Bagos *et al.*, 2004b)
Outer membrane lipoprotein	LIPO	http://services.cbu.uib.no/tools/lipo	(Berven *et al.*, 2006)
	LIPOP v. 1.0	http://www.cbs.dtu.dk/services/LipoP/	(Juncker *et al.*, 2003)

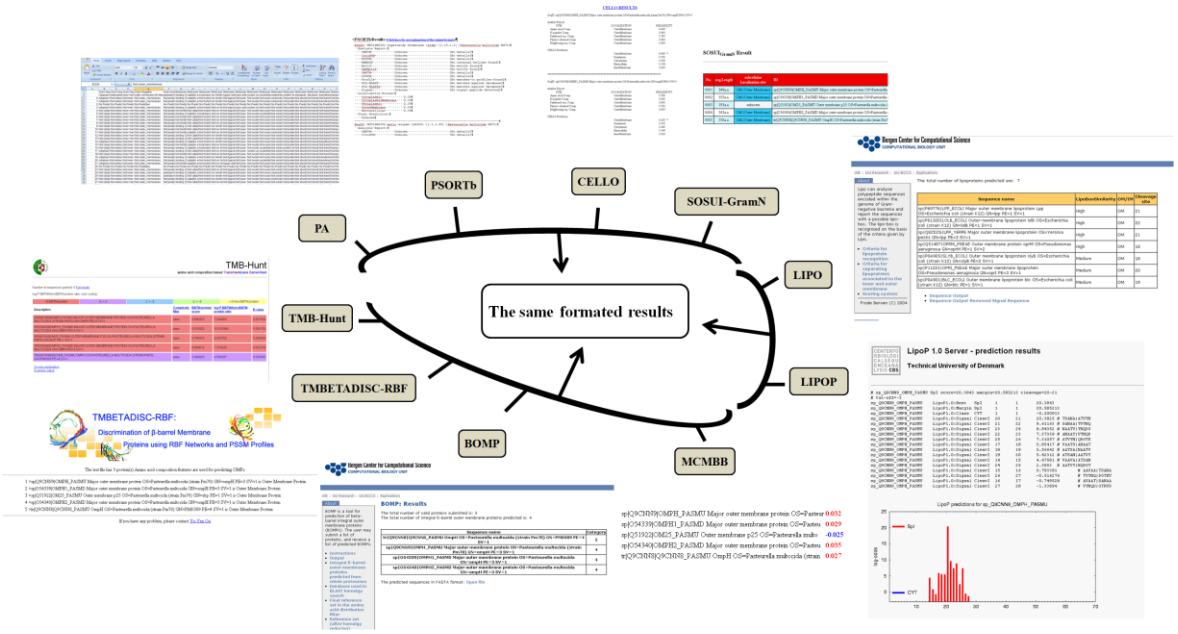


Figure 2-3. Diagram showing that different predictors give various formats of results. Therefore, in-house built perl scripts were used to change these results into comparable formats.

predictors. For the lipoprotein predictors, the criteria were varied from positive predicted proteins obtained by at least one or two predictors. These criteria were evaluated by the calculation of accuracy, recall/sensitivity, specificity and Matthews correlation coefficient (MCC) from the set of *P. multocida* proteins with known subcellular localization in the Uniprot protein database. The criteria that maximized the above parameters were selected as optimal and used for results filtering. Subsequently, consensus predictions from the three groups were combined, representing a single list of predicted OMPs. The formulas are shown below where TP represents the number of true positives (OMPs predicted as OMPs), TN the true negatives (non-OMPs predicted as non-OMPs), FP the false positives (non-OMPs predicted as OMPs) and FN the false negatives (OMPs predicted as non-OMPs). This optimization tended to reduce most of the false positive and retain most of the true positive proteins.

$$Accuracy\,(\%\,of\ correct\ predictions) \;=\; \frac{(TP+TN)}{(TP+FP+TN+FN)}$$

$$\mathrm{Re}call\ or\ Sensitivity\,(\%\,of\ true\ positive\ identification) \;=\; \frac{TP}{(TP+FN)}$$

$$Specificity\,(\%\,of\ true\ negative\ identification) \;=\; \frac{TN}{(TN+FP)}$$

$$MCC\,(Measurement\ of\ prediction\ quality) \;=\; \frac{(TP\times TN - FP\times FN)}{\sqrt{(TP+FN)(TP+FP)(TN+FP)(TN+FN)}}$$

2.2.5 Consensus prediction and integration

Once the optimum criteria were selected, the predicted proteins from these three groups of predictors were filtered. Proteins that could pass the criteria threshold of each predictor group were integrated into a single list of predicted OMPs.

2.2.6 Manual confirmation and functional prediction of the predicted proteins

After the consensus prediction, each predicted protein was manually confirmed as being outer membrane-associated by using public database searches and sequence analyses. The

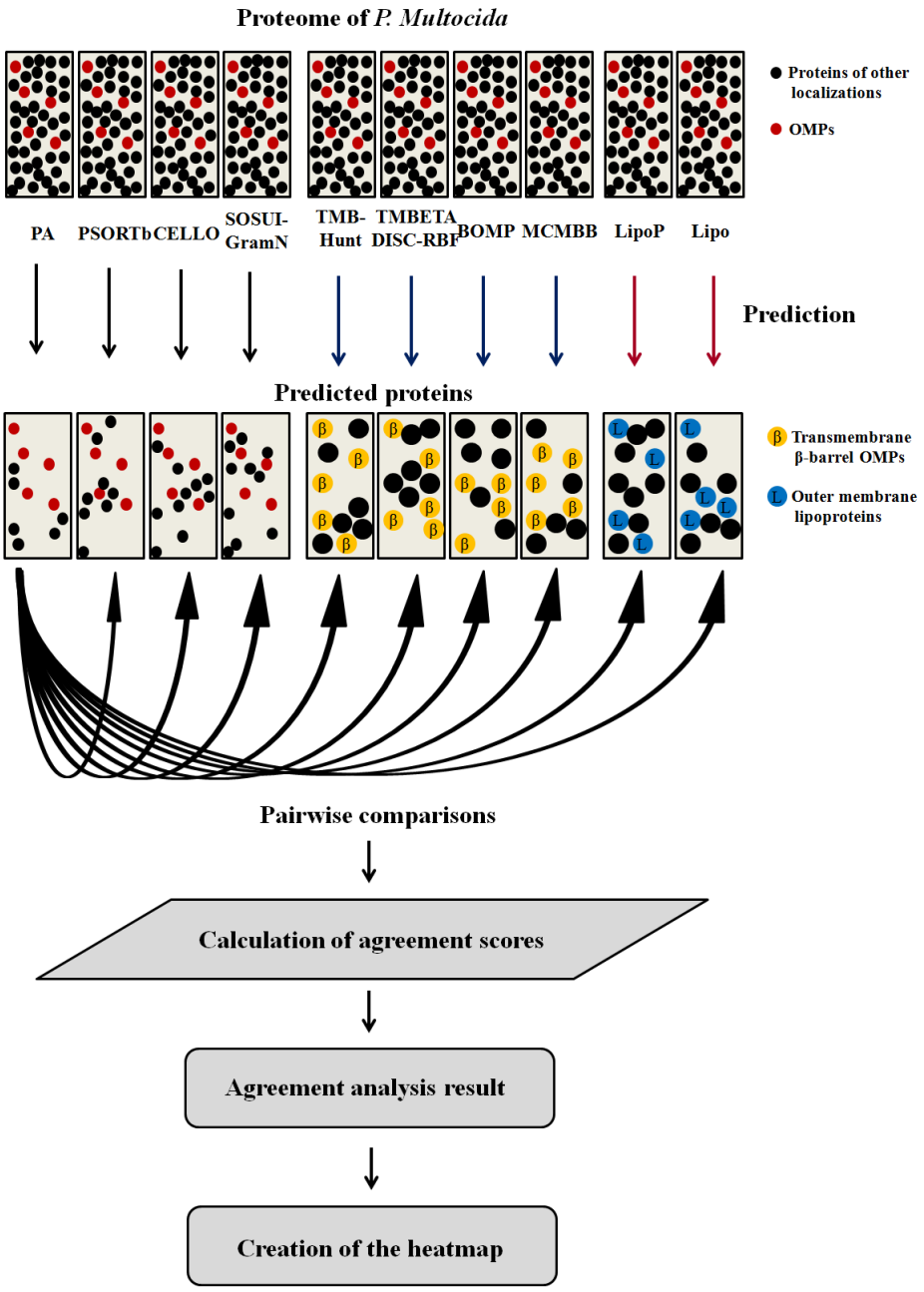


Figure 2-4. This diagram summarizes the procress of the agreement analysis. Ten predictors were used to predict OMPs from the avian strain genome of *P. multocida*. The predicted proteins from pairs of predictors were compared and the agreement scores for each predictor pair were calculated by in-house built perl scripts. The result was represented by heatmap plotting. The black circles show possible false positive results during the prediction step.

PubMed database (http://www.ncbi.nlm.nih.gov/pubmed) was used to retrieve published experimental information relevant to the predicted proteins. The UniProt protein database (http://www.uniprot.org/) was searched for homology and domain/motif, protein-protein interactions, and functional and structural predictions. Structural homology was examined by using the HHpred program (http://toolkit.tuebingen.mpg.de/hhpred, (Söding, 2005)). The STRING protein interaction database (http://string-db.org/) was used to identify whether the predicted proteins interacted with any known proteins or were members of any characterized pathways. Taken together, these analyses confirmed the proteins as confidently predicted putative OMP candidates.

2.2.7 Analysis of proteins that were filtered out

The use of filtering criteria aims to reduce the number of false positive proteins; however it may open a probability of losing certain true positive proteins. Therefore, predicted proteins that were filtered out were re-examined by the manual confirmation step to identify possible left-out OMPs as described in **section 2.2.6 (Figure 2-5)**.

2.2.8 Physicochemical properties of the predicted OMPs

Physicochemical properties, e.g. molecular weight, length of protein sequence, theoretical pI, grand average of hydropathicity (GRAVY) score, aliphatic index, charge, number of β-strands and helices, of the putative OMPs were predicted by the ProtParam program (http://expasy.org/tools/protparam.html), TMBETA-NET (Gromiha & Suwa, 2006a) and TMHMM (Krogh *et al.*, 2001).

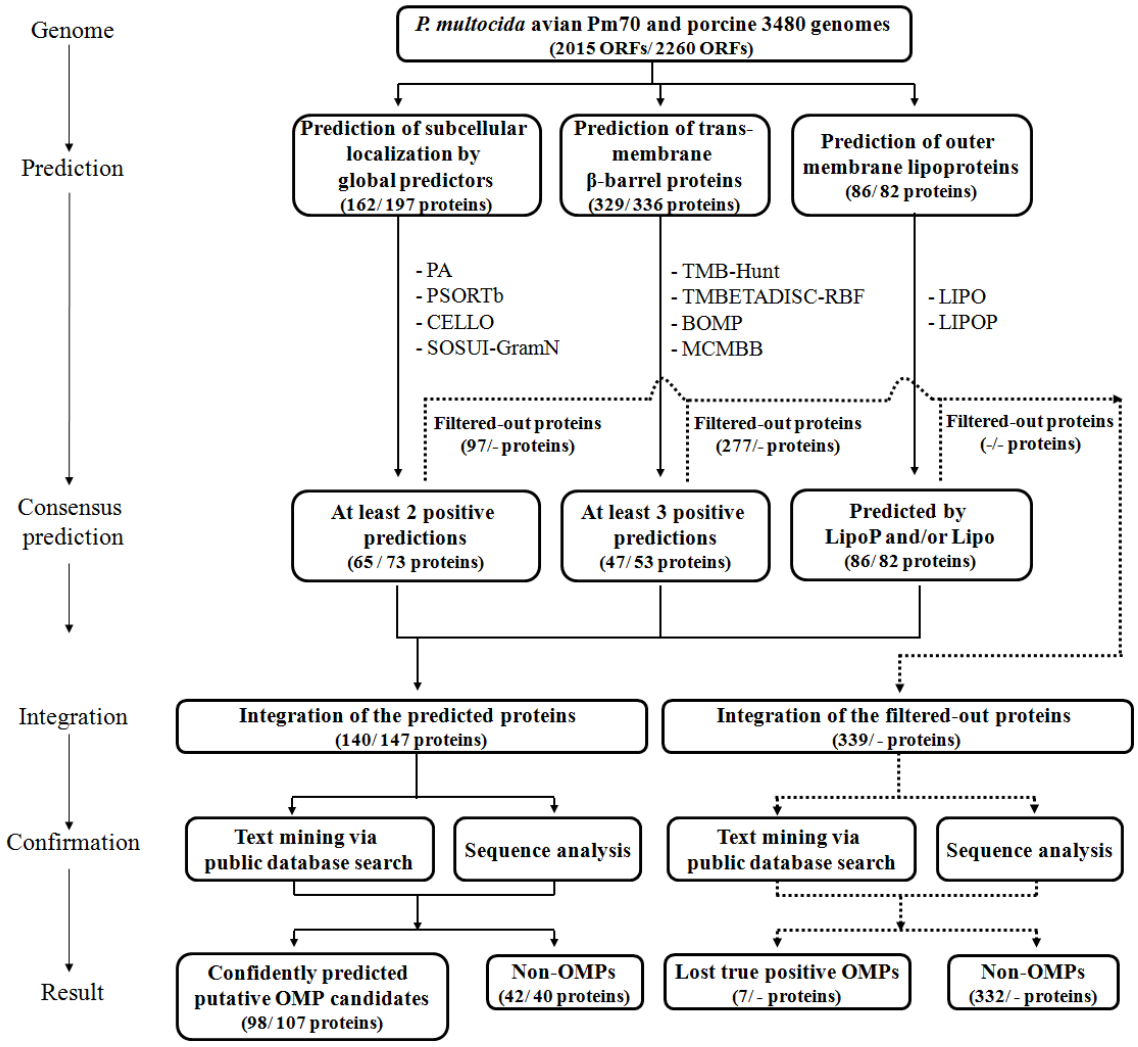


Figure 2-5. Modified bioinformatic prediction workflow (from **Figure 2-2**) for the examination of filtered-out predicted proteins to identify possible left-out putative OMPs. This is indicated by the dashed lines. Filtered-out proteins from each predictor group were integrated and processed through the manual confirmation step to identify any true positive OMPs lost during the consensus prediction step. The numbers of predicted proteins in each step are shown in parentheses: the first number represents proteins from the avian strain genome and the second number from the porcine strain genome.

2.3 Results

2.3.1 Prediction of OMPs using different predictors

Outer membrane proteins were predicted, by ten different bioinformatic programs **(Table 2-2)**, from the two available genomes of *P. multocida*; the genome of avian strain Pm70 and the genome of porcine strain 3480. These programs were categorized into three groups: subcellular localization predictors (PA, PSORTb, CELLO, SOSUI-GramN), transmembrane β-barrel protein predictors (TMB-Hunt, TMBETADISC-RBF, BOMP, MCMBB), and lipoprotein predictors (LIPO and LIPOP). Individual programs predicted different numbers of proteins. The use of these ten predictors in combination predicted 421 putative OMPs from the avian strain genome (20.9% of the genome) and 439 proteins from the porcine strain genome (19.4% of the genome) **(Figure 2-2, Appendix Tables 2-1 and 2-2)**.

The subcellular localization predictors identified 162 putative OMPs from the avian strain genome and 197 proteins from the porcine strain genome **(Figure 2-6A)**. CELLO identified the highest (91 and 108) and PSORTb identified the lowest (49 and 63) number of predicted proteins from the avian and porcine strain genomes, respectively. For the avian strain genome, 97 proteins were predicted by only a single program: 35, 24, 3 and 35 by CELLO, PA, PSORTb and SOSUI-GramN, respectively. Similarly, 124 proteins were identified by a single predictor from the porcine strain genome: 50, 30, 5 and 39 by CELLO, PA, PSORTb and SOSUI-GramN, respectively. Twenty-four proteins were identified from the avian strain genome and 22 from the porcine strain genome using all four programs. The use of two or three programs predicted a total of 41 proteins from the avian strain genome and 51 proteins from the porcine strain genome.

The transmembrane β-barrel protein predictors identified 329 putative β-barrel proteins from the avian strain genome and 336 proteins from the porcine strain genome **(Figure 2-**

6B). TMB-Hunt identified the highest number of predicted proteins (168) from the avian strain genome, while MCMBB identified the highest number of predicted proteins (184) from the porcine strain genome. BOMP identified the lowest number of predicted proteins (40 and 48) from the avian and porcine strain genomes, respectively. For the avian strain genome, 231 proteins were predicted by only a single program: 70, 113, 43 and 5 proteins by MCMBB, TMB-Hunt, TMBETADISC-RBF and BOMP, respectively. Similarly, 225 proteins were identified by only a single predictor from the porcine strain genome: 84, 84, 46 and 11 proteins by MCMBB, TMB-Hunt, TMBETADISC-RBF and BOMP, respectively. Nineteen proteins were predicted by all four programs in both the avian and porcine strain genomes. The use of two or three programs predicted a total of 79 proteins from the avian strain genome and 92 proteins from the porcine strain genome.

The lipoprotein predictors identified 86 proteins from the avian strain genome and 82 proteins from the porcine strain genome **(Figure 2-6C)**. LIPO predicted 73 proteins from the avian strain genome and 75 from the porcine strain genome whereas LIPOP predicted 69 proteins from the avian strain genome and 67 from the porcine strain genome. Together, LIPO and LIPOP predicted 56 and 60 proteins from the avian and porcine strain genomes, respectively. However, LIPO identified 17 unique lipoproteins from the avian strain genome and 15 from the porcine strain genome, whereas LIPOP identified 13 unique lipoproteins from the avian strain genome and seven from the porcine strain genome.

Comparison of the predicted OMPs by the three groups of predictors revealed that the use of one group of predictors alone identified 290 proteins from the avian strain genome and 283 proteins from the porcine strain genome, whereas a combination of two groups of predictors identified 106 proteins from the avian strain genome and 130 proteins from the porcine strain genome **(Figure 2-7)**. The use of all three groups of predictors identified 25 proteins from the avian strain genome and 24 proteins from the porcine strain genomes.

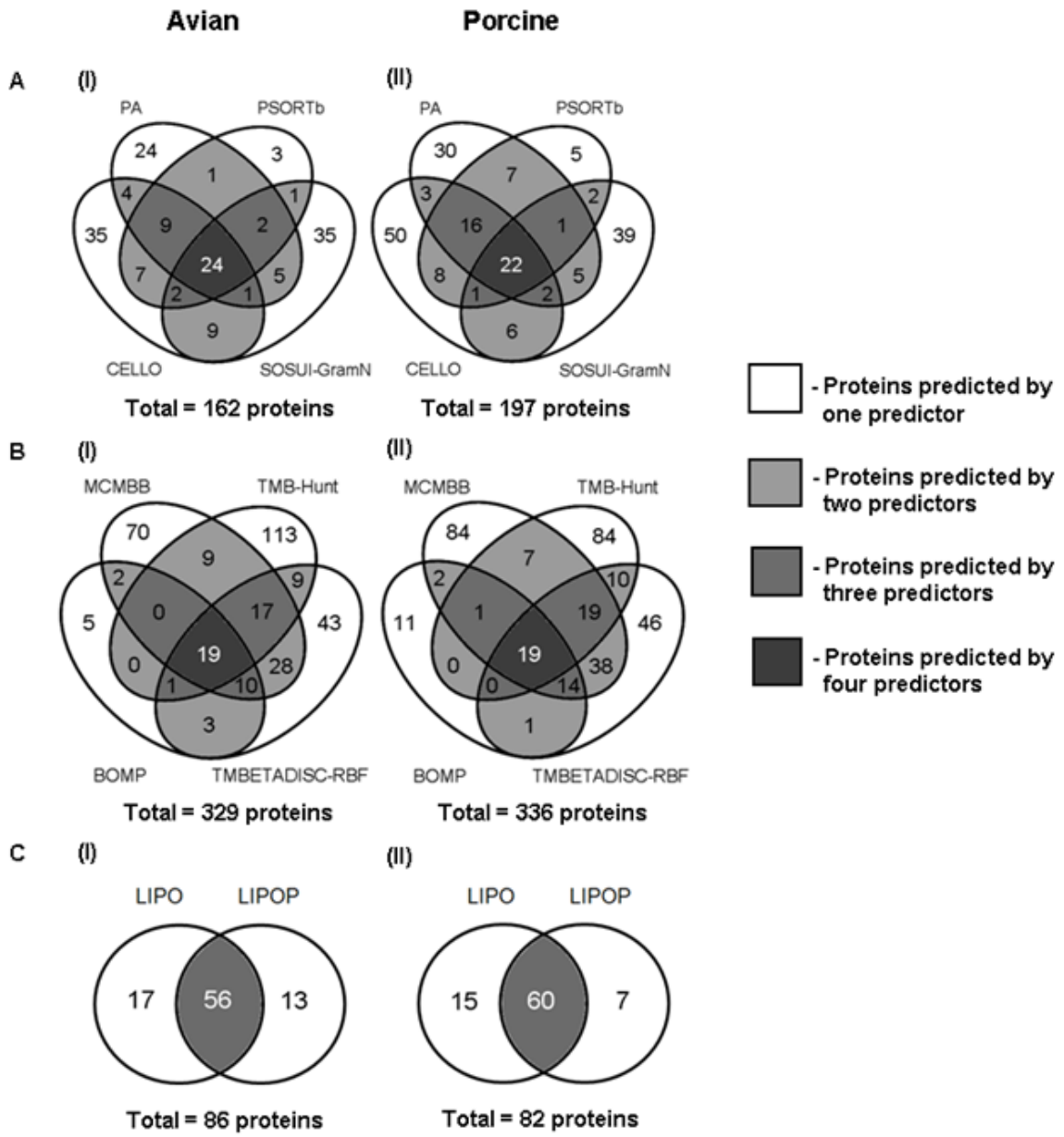


Figure 2-6. Diagrams showing within-group comparisons of the numbers of proteins predicted by three groups of predictors: subcellular localization (A), transmembrane β-barrel protein (B), lipoprotein (C) predictors. The diagrams on the left side represent the avian strain genome and those on the right the porcine strain genome. Indicated are the numbers of proteins predicted by one, two, three or four predictors.

Noticeably, the transmembrane β-barrel protein predictors predicted a high number of proteins (217 and 202) that were not predicted by the other two groups of predictors.

2.3.2 Agreement between pairs of predictors

The analysis in **Figure 2-8** shows different degrees of agreement between pairs of outer membrane predictors. For the subcellular localization predictors, prediction by pairs of PA and PSORTb as well as PSORTb and CELLO resulted in high agreement scores (0.74 and 0.86, respectively). Pairing of PSORTb with TMBETADISC-RBF and MCMBB also produced high agreement scores (0.90 and 0.76, respectively). For the transmembrane β-barrel protein predictors, predictions by pairing of BOMP with MCMBB and TMBETADISC-RBF as well as MCMBB with TMBETADISC-RBF showed moderate scores (0.57 in average), while pairs of LIPO and LIPOP had a higher agreement score of 0.77 for lipoprotein prediction. The disagreement between lipoprotein predictors and the others was clearly shown with scores of less than 0.5. Subcellular localization predictors discriminate between proteins belonging to different locations. Although these predictors predict a wide range of outer membrane-located proteins, and some of these predictors incorporate the prediction of transmembrane β-barrel proteins and lipoproteins as parts of their programs, some OMPs were possibly mispredicted or excluded, as confirmed by the low agreement score between subcellular localization and lipoprotein predictors. Conversely, the transmembrane β-barrel and lipoprotein predictors differentiate between specific groups of OMPs; they are unable to predict all outer membrane-localized proteins. Therefore, a combination of the subcellular localization, transmembrane β-barrel and lipoprotein predictors resulted in better coverage of the predicted OMPs than the use of a single predictor or group of predictors.

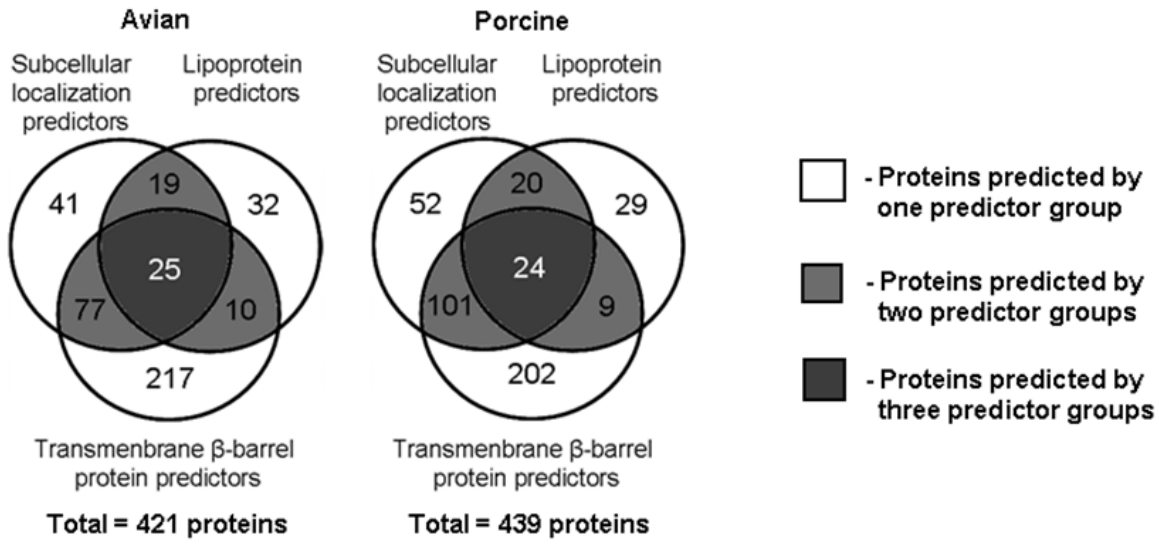


Figure 2-7. Diagrams showing between-group comparisons of the numbers of proteins predicted by three groups of predictors: subcellular localization, transmembrane β-barrel protein, lipoprotein predictors. The diagram on the left side represents the avian strain genome and that on the right the porcine strain genome. Indicated are the numbers of proteins predicted by one, two or three predictor groups.

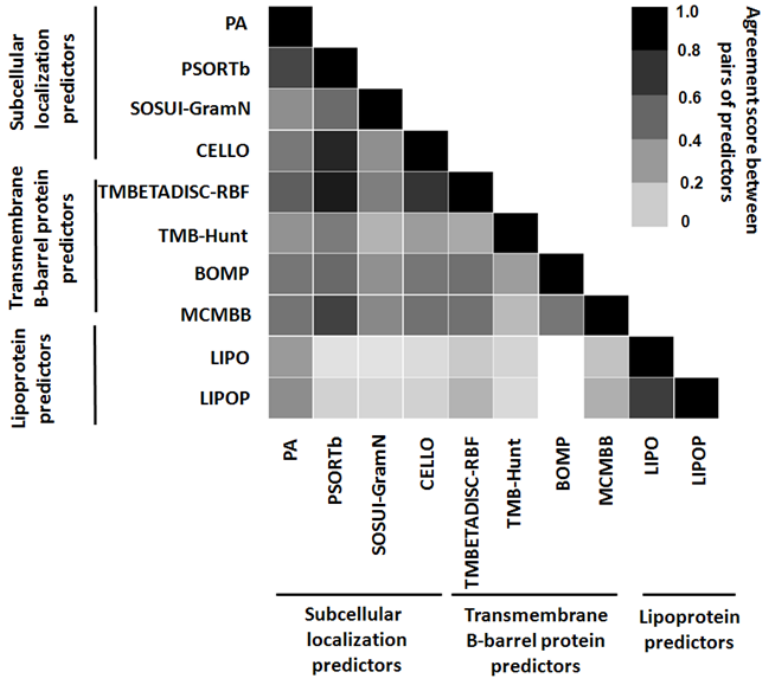


Figure 2-8. Analysis of agreement between pairs of different bioinformatic programs (10 programs classified into three groups: subcellular localization, transmembrane β-barrel and lipoprotein predictors) used for the prediction of OMPs within the avian strain genome. Each square represents the color coded agreement score which corresponds to the proportion of commonly predicted proteins for pairs of predictors. The agreement score ranges from 0 for the lowest agreement (white) to 1 for the highest (black).

2.3.3 Consensus predicted OMPs

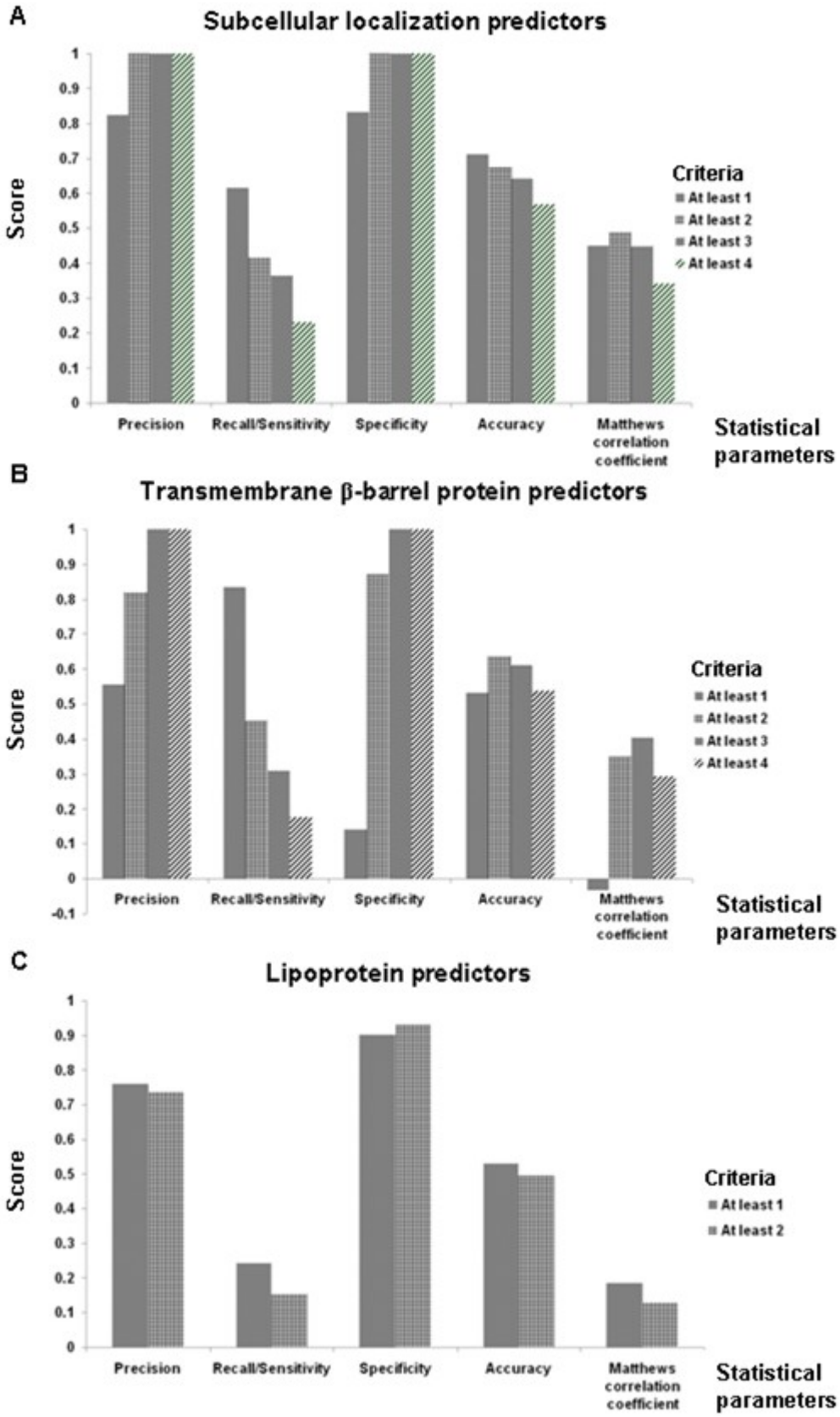
The above analyses indicated different levels of agreement between pairs of predictors. The use of multiple predictors for subcellular localization, transmembrane β-barrel and lipoprotein predictions produced a large number of predicted proteins many of which are potential false positives. Therefore, the predicted results from individual predictors in each group were filtered using various criteria. The predicted *P. multocida* proteins of known localizations derived from the Uniprot database were used as a training data set in the measurement of precision, recall, specificity, accuracy and Mathews correlation coefficient (MCC) (Ou *et al.*, 2008). The criteria which gave the highest scores of these five parameters were selected for the consensus prediction. For the subcellular localization predictors **(Figure 2-9A)**, prediction by at least two predictors was selected because this threshold gave the highest precision, specificity and MCC score. With this selected threshold, all the false positives were removed. Similarly, prediction by at least three predictors was chosen for the transmembrane β-barrel protein predictors **(Figure 2-9B)** because this threshold gave the highest precision, specificity and MCC scores. In both cases, increased precision and specificity occurred as the number of predictors increased whereas, conversely, recall and accuracy decreased. For the lipoprotein predictors **(Figure 2-9C)**, prediction by at least one predictor was selected as this resulted in the highest precision, recall, accuracy and MCC scores.

The proteins predicted by each group of predictors were filtered using these optimized criteria and resulted in 65 consensus predicted proteins from the avian strain genome and 73 proteins from the porcine strain genome for the subcellular localization predictors; 47 and 53 proteins from the avian and porcine strain genomes, respectively, for the β-barrel transmembrane protein predictors; and 86 and 82 proteins from the avian and porcine strain genomes, respectively, for the lipoprotein predictors **(Figure 2-2).** Integration of the consensus-predicted proteins from these three groups subsequently yielded 140 proteins

from the avian strain genome and 147 proteins from the porcine strain genome **(Figure 2-2)**. Of the 140 proteins predicted from the avian strain genome and 147 proteins from the porcine strain genome at the integration step, 27 proteins from the avian strain genome and 24 proteins from the porcine strain genome were filtered out by the consensus threshold of the subcellular localization predictor group but not by the consensus threshold of the β-barrel transmembrane protein and/or that of the lipoprotein predictor groups. Similarly, 36 proteins from the avain strain genome and 34 proteins from the porcine strain genome were filtered out by the consensus threshold of the β-barrel transmembrane protein predictor group but not by the consensus threshold of the subcellular localization and/or that of the lipoprotein predictor groups. No proteins from either genome were removed from the lipoprotein predictor group by its consensus threshold.

2.3.4 Manual curation of the predicted proteins

In the final step, published information available on the predicted proteins was searched, using text mining and sequence analysis, to confirm their outer membrane location. Forty-two proteins (30%) from the avian strain genome and 40 proteins (27%) from the porcine strain genome were removed at the manual confirmation stage. Thirty-one of these proteins were identified in both genomes and included 19 cytoplasmic or inner membrane proteins, 11 periplasmic proteins, two secreted proteins and one phage protein. In this way, 98 proteins from the avian strain genome and 107 proteins from the porcine strain genome were confirmed as being confidently-predicted OMPs **(Figure 2-2)**. These proteins accounted for 4.9% of the avian strain genome and 4.7% of the porcine strain genome. Details of the confidently predicted OMPs from the avian strain genome are given in **Table 2-3**. Eighty-nine (91%) of the predicted OMPs in the avian strain genome were also detected in the porcine strain genome, indicating that these two outer membrane proteomes are very similar. Eighteen (17%) of the predicted OMPs from the porcine strain


A
Subcellular localization predictors
Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Criteria
At least 1
At least 2
At least 3
At least 4
Precision
Recall/Sensitivity
Specificity
Accuracy
Matthews correlation coefficient
Statistical parameters
B
Transmembrane β-barrel protein predictors
Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
Criteria
At least 1
At least 2
At least 3
At least 4
Precision
Recall/Sensitivity
Specificity
Accuracy
Matthews correlation coefficient
Statistical parameters
C
Lipoprotein predictors
Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Criteria
At least 1
At least 2
Precision
Recall/Sensitivity
Specificity
Accuracy
Matthews correlation coefficient
Statistical parameters

genome had no homologous proteins in the avian strain genome; most of these were hypothetical proteins. However, these proteins included an Omp100 adhesin/invasin homologue in *Aggregatibacter aphrophilus* and two uncharacterized TonB-dependent receptors.

Of the 98 confidently predicted OMPs of the avian strain genome, 59 (60%) were predicted by subcellular localization, 44 (45%) by transmembrane β-barrel, and 49 (50%) by lipoprotein predictors **(Figure 2-10)**. Thirty-one proteins were identified as transmembrane β-barrel proteins by both subcellular localization and transmembrane β-barrel predictors. A further five were identified as transmembrane β-barrel proteins by the β-barrel predictors alone; two of these were hypothetical β-barrel proteins (PM0519 and PM1772) which might have novel functions. Thirty-two proteins were uniquely predicted to be outer membrane lipoproteins which were consistent with the agreement analysis. However, almost half of these were of unknown function. A further nine proteins were identified as lipoproteins by both lipoprotein and subcellular localization predictors. Thirteen proteins were identified only by the subcellular localization predictors. Four of these (OmpW and the TonB-dependent receptors PM0745, PM1081 and PM1428) contain transmembrane β-barrel domains but were filtered out by the transmembrane β-barrel predictors since they did not pass the criteria. Two proteins were identified by both transmembrane β-barrel and lipoprotein predictors and six proteins by all three groups of predictors.

Table 2-3. Confidently-predicted putative OMPs identified from the genome of avian *P. multocida* strain Pm70 by 10 predictors, categorized into three groups (subcellular localization, transmembrane β-barrel and lipoprotein predictors) and subjected to the bioinformatic process described in **Figure 2-2**. OMPs predicted from the porcine *P. multocida* genome strain 3480 were compared. Physicochemical properties including molecular weight (MW), PI, aliphatic index, GRAVY score, number of transmembrane helices and β-strands, length and charge of each protein are also shown. Proteins were classified by predictor groups.

No	ID	Name	Presence in the porcine genome[a]	Bioinformatic prediction of the OMPs[b]										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
1. Proteins predicted by subcellular localization and transmembrane β-barrel and predictors (31)																					
1	PM0040	PfhR	√	+	+	-	+	+	+	+	+	-	-	81.332	9.41	69.7	-0.646	0	31	727	36
2	PM0056	LspB_1	√	+	+	+	+	+	+	+	+	-	-	54.024	9.34	85.55	-0.359	0	22	576	16
3	PM0058	LspB_2	√	+	+	+	+	+	+	+	+	-	-	52.7357	9.6	85.17	-0.336	0	21	573	22
4	PM0076	EstA	√	-	+	-	+	+	-	+	+	-	-	74.5689	7.72	82.3	-0.285	1	32	679	9
5	PM0300	HgbA	√	+	+	+	+	+	+	+	+	-	-	109.713	8.92	70.79	-0.666	0	34	963	23.5
6	PM0336	HgbB	√	+	+	+	+	+	+	+	-	-	-	113.843	9.1	69.4	-0.705	1	40	989	33
7	PM0337	HgbB	√	+	+	+	+	+	+	+	+	-	-	113.34	8.84	73.36	-0.625	1	40	997	24.5

[a] OMPs predicted from the genome of porcine *P. multocida* strain 3480; '√' = positive prediction and '-' = negative prediction
[b] Bioinformatic prediction; '+' = positive prediction and '-' = negative prediction

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome[a]	Bioinformatic prediction of the OMPs[b]										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Alphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
8	PM0388	OmpH_1	√	+	+	+	+	+	+	+	+	-	-	37.4502	8.82	82.36	-0.276	0	16	348	6
9	PM0389	OmpH_2	√	+	+	+	+	+	-	+	+	-	-	38.7777	8.4	84.6	-0.292	0	15	350	6
10	PM0663	NanH	√	-	+	-	+	+	-	+	+	-	-	93.3456	8.53	81.24	-0.434	3	44	832	12.5
11	PM0714	Hsf_1	√	+	-	+	-	+	+	-	+	-	-	276.155	5.38	80.03	-0.388	3	25	2712	-27.5
12	PM0741	HmbR	√	+	+	-	+	+	+	+	+	-	-	89.5434	9.19	72.67	-0.608	0	35	784	25
13	PM0786	OmpA	√	+	+	-	+	+	-	+	+	-	-	38.0309	9.09	82.72	-0.226	0	14	353	10.5
14	PM0803	TonB-dependent receptor	√	+	+	+	+	+	+	+	+	-	-	90.9927	8.58	68.46	-0.672	2	30	792	13.5
15	PM0831	OmpH_3	√	+	+	-	+	+	+	+	+	-	-	34.9712	8.35	80.03	-0.368	0	14	313	5.5
16	PM0852	RcpA	√	+	+	+	+	+	+	-	+	-	-	51.1167	6.21	94.32	-0.152	1	19	470	0
17	PM0853	RcpC	√	-	-	+	+	+	+	-	+	-	-	30.1015	6.05	108.26	-0.186	0	13	270	-0.5
18	PM0998	MipA/OmpV family protein	√	+	+	+	+	+	+	+	+	-	-	28.623	9.63	82.76	-0.151	0	12	257	17.5

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
19	PM1000	NanB	√	-	-	+	+	+	+	+	+	-	-	121.45	9.02	77.71	-0.557	0	39	1080	26.5
20	PM1025	Opa	√	+	+	+	+	+	+	+	+	-	-	20.5124	9.39	81.24	-0.046	1	12	186	8
21	PM1069	FadL	√	+	+	+	+	+	+	+	+	-	-	46.0612	9.12	82.87	-0.1	0	16	428	11
22	PM1282	OM hemin receptor	√	+	+	+	+	+	-	+	+	-	-	88.1373	8.92	83.43	-0.461	0	34	778	21.5
23	PM1426	Phospholipase A/OmpLA	√	+	+	+	+	+	-	+	+	-	-	35.5805	9.39	82.22	-0.521	0	13	306	14.5
24	PM1515	Conserved hypothetical protein	√	-	-	+	+	+	-	+	+	-	-	58.7064	9.61	75.91	-0.56	0	23	509	22.5
25	PM1543	Hypothetical protein	√	+	+	+	+	+	+	+	+	-	-	26.8803	9.12	83.05	-0.394	0	11	236	8.5
26	PM1570	Hsf_2	√	+	-	+	-	+	+	+	+	-	-	130.963	7.53	78.14	-0.268	2	22	1299	4.5
27	PM1600	LptD/Imp/OstA	√	+	+	+	+	+	-	+	+	-	-	90.6472	9.03	73.63	-0.708	0	36	782	20

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Alphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
28	PM1622	HasR	√	+	+	+	+	+	+	+	+	-	-	95.9098	9.11	78.09	-0.478	0	32	848	23.5
29	PM1717	OM autotransporter	√	+	+	+	-	+	+	+	+	-	-	95.1859	8.43	77.47	-0.546	0	42	850	23
30	PM1809	Omp85 family protein/YtfM	√	-	+	-	+	+	-	+	+	-	-	67.106	9.29	85.02	-0.46	0	28	586	17.5
31	PM1992	Oma87	√	+	+	+	+	+	+	+	+	-	-	87.7612	6.3	79.75	-0.39	3	37	791	0
2. Proteins predicted only by transmembrane β-barrel predictors (5)																					
1	PM0266	Mce/PqiB	√	-	-	-	-	+	+	-	+	-	-	96.9674	6.76	102.25	-0.105	3	40	884	8
2	PM0395	YccT	√	-	-	-	-	+	+	-	+	-	-	23.742	9.2	88.17	-0.201	0	9	218	4.5
3	PM0519	Conserved hypothetical protein	√	-	-	-	-	+	+	-	+	-	-	12.6203	9.58	82.19	-0.441	0	9	114	6.5
4	PM1772	Hypothetical protein	√	-	-	-	-	+	+	-	+	-	-	7.11717	10.09	88.62	-0.358	0	6	65	6

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
5	PM1808	OmpL41/YtfN-like protein	√	-	-	-	-	+	-	+	+	-	-	142.064	5.87	107.28	-0.113	3	38	1300	-5
3. Proteins predicted only by lipoprotein predictors (32)																					
1	PM0016	Lipoprotein	√	-	-	-	-	-	-	-	-	+	+	11.4983	9.48	92.5	-0.497	0	5	100	6.5
2	PM0072	Lipoprotein	√	-	-	-	-	-	-	-	-	+	+	13.1421	9.04	100.09	-0.247	0	7	116	6
3	PM0246	LolB	√	-	-	-	-	-	-	-	-	+	+	24.3777	8.55	85.75	-0.581	0	11	207	6
4	PM0442	Lipoprotein	√	-	-	-	-	-	-	-	-	+	+	23.9973	5.04	72.27	-0.392	0	10	229	-9
5	PM0513	MltB	√	-	-	-	-	-	-	-	-	+	+	41.5161	9.53	84.42	-0.448	0	18	364	14.5
6	PM0627	Lipoprotein NlpC/P60	√	-	-	-	-	-	-	-	-	+	+	18.161	9.54	86.42	-0.174	0	10	159	11
7	PM0708	Slp	√	-	-	-	-	-	-	-	-	-	+	20.575	8.91	100.78	-0.209	0	9	179	5
8	PM0758	Lipoprotein	√	-	-	-	-	-	-	-	-	+	+	29.3199	8.68	96.07	-0.195	1	15	267	7.5
9	PM0931	LppA	√	-	-	-	-	-	-	-	-	-	+	18.4605	7.85	116.01	-0.017	0	6	163	3
10	PM0982	Lipoprotein	√	-	-	-	-	-	-	-	-	+	+	30.0912	4.68	109.46	-0.032	0	15	261	-11.5

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
11	PM1002	Lipoprotein	√	-	-	-	-	-	-	-	-	+	-	43.520	8.36	101.56	-0.096	0	-	391	5.5
12	PM1044	Lipoprotein	√	-	-	-	-	-	-	-	-	+	+	28.8309	9.04	70.17	-1	0	10	242	10
13	PM1050	NlpB	√	-	-	-	-	-	-	-	-	+	+	37.3733	7.7	87.45	-0.372	0	17	337	2
14	PM1060	Conserved hypothetical protein	√	-	-	-	-	-	-	-	-	+	+	8.24469	4.63	116.16	0.4	1	3	73	-3
15	PM1064	Lipoprotein E/OmpP4	√	-	-	-	-	-	-	-	-	+	+	30.119	9.38	70.33	-0.49	0	15	272	10.5
16	PM1073	Lipoprotein	√	-	-	-	-	-	-	-	-	+	+	21.4116	5.69	87.01	-0.388	0	7	184	-1.5
17	PM1077	HlpB	√	-	-	-	-	-	-	-	-	-	+	20.5573	5.21	86.58	-0.601	0	9	184	-4
18	PM1190	Peptidase M48B family protein	√	-	-	-	-	-	-	-	-	-	+	28.1523	9.42	79.11	-0.277	1	12	257	11
19	PM1215	RlpB	√	-	-	-	-	-	-	-	-	+	+	18.8881	9.33	101.68	-0.086	0	8	167	6

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
20	PM1501	VacJ	√	-	-	-	-	-	-	-	-	+	+	27.5284	7.71	86.38	-0.293	1	7	246	1.5
21	PM1514	PlpE	√	-	-	-	-	-	-	-	-	+	+	37.2715	6.56	73.29	-0.714	0	13	331	3
22	PM1518	PlpP	√	-	-	-	-	-	-	-	-	+	+	37.3624	6.03	65.43	-0.757	0	11	348	-0.5
23	PM1578	Lipoprotein	√	-	-	-	-	-	-	-	-	-	+	35.8607	6.76	88.76	-0.057	1	15	339	1
24	PM1614	LppB/NlpD	√	-	-	-	-	-	-	-	-	+	+	49.7348	9.19	72.81	-0.407	0	14	467	11.5
25	PM1720	ComL	√	-	-	-	-	-	-	-	-	+	+	29.3501	7.74	82.38	-0.352	0	11	260	2.5
26	PM1730	PlpA/MetQ	√	-	-	-	-	-	-	-	-	+	+	30.2324	5.2	93.66	-0.255	1	12	276	-5
27	PM1798	Lipoprotein	√	-	-	-	-	-	-	-	-	+	+	19.2247	6.4	88.68	-0.327	0	6	171	1
28	PM1805	HlpB	√	-	-	-	-	-	-	-	-	-	+	19.4394	6.06	100.4	-0.066	0	4	177	0
29	PM1827	Hypothetical protein	√	-	-	-	-	-	-	-	-	+	+	17.452	9.5	80.62	-0.637	0	7	160	10.5
30	PM1886	SmpA	√	-	-	-	-	-	-	-	-	+	+	15.5339	7.81	94.6	-0.12	0	8	137	2
31	PM1939	Lipoprotein		-	-	-	-	-	-	-	-	+	-	16.700	6.51	76.23	-0.351	0	-	151	1.5
32	PM2008	PilW/PilF	√	-	-	-	-	-	-	-	-	+	+	20.9114	7.77	72.93	-0.534	0	7	181	6

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
4. Proteins predicted by subcellular localization and lipoprotein predictors (9)																					
1	PM0554	Lpp/Pcp	√	+	-	+	-	-	-	-	-	+	+	15.5879	9.07	107.53	0.284	2	11	154	3.5
2	PM0586	Plp4	√	+	+	+	+	-	-	-	-	+	+	30.0749	8.99	74.19	-0.422	0	12	272	8
3	PM0659	Lipoprotein	√	-	+	-	+	-	-	-	-	-	+	214.428	5.93	88.88	-0.347	1	38	1905	-6
4	PM0846	TadD	√	-	-	+	+	-	-	-	-	+	+	28.7109	9.45	105.53	-0.175	1	14	257	10.5
5	PM0966	Pal/Omp P6	√	+	+	-	+	-	-	-	-	+	+	16.2132	7.79	80.67	-0.293	0	8	150	2.5
6	PM1321	MltC	√	+	-	+	+	-	-	-	-	+	+	40.2941	9.63	90.73	-0.229	1	23	358	17.5
7	PM1444	GlpQ	√	+	+	-	-	-	-	-	-	+	+	41.1622	6.57	83.1	-0.436	0	18	373	3
8	PM1826	Hypothetical protein	√	+	-	+	-	-	-	-	-	+	+	25.5931	9.78	79.34	-0.333	2	11	243	15
9	PM1980	IbeB	√	+	+	-	+	-	-	-	-	+	+	51.9452	8.64	98.98	-0.279	1	20	463	8
5. Proteins predicted only by subcellular localization predictors (13)																					
1	PM0015	Hypothetical protein	√	-	-	+	+	-	-	-	-	-	-	23.5143	5.61	80.48	-0.474	0	11	207	-2.5

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
2	PM0234	Hypothetical protein	√	-	-	+	+	-	-	-	-	-	-	81.3928	5.79	86.99	-0.544	0	32	708	-4.5
3	PM0243	NlpD-like protein	√	-	-	+	+	-	-	-	-	-	-	59.5582	9.07	83.2	-0.564	0	22	531	14
4	PM0331	OmpW	√	+	+	+	-	-	-	-	-	-	-	21.8701	9.16	96.08	0.175	0	10	204	6.5
5	PM0698	Mod_2	-	-	+	-	+	-	-	-	-	-	-	72.5971	5.29	91.98	-0.457	1	27	636	-11.5
6	PM0745	TonB-dependent receptor	√	+	-	-	+	-	-	-	-	-	-	105.082	7.54	80.34	-0.533	2	32	925	21.5
7	PM1081	TonB-dependent receptor	√	+	+	+	+	-	-	-	-	-	-	90.9033	8.87	79.57	-0.513	1	38	809	30.5
8	PM1225	ComE/PilQ	√	+	+	+	+	-	-	-	-	-	-	49.2065	7.79	101.24	-0.197	0	22	444	8
9	PM1428	TonB-dependent receptor	√	+	+	+	+	-	-	-	-	-	-	90.8504	8.98	77.38	-0.521	2	34	805	30
10	PM1819	Virulence factor SrfB	√	-	+	+	-	-	-	-	-	-	-	117.436	5.56	89.01	-0.426	1	34	1024	-13
11	PM1926	RlpA-like protein	√	+	-	-	+	-	-	-	-	-	-	33.1802	9.79	92.89	-0.374	0	18	295	23

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
12	PM1993	Skp/Outer membrane p25	√	+	+	-	+	-	-	-	-	-	-	21.4084	8.55	89.59	-0.501	0	10	193	3
13	PM2009	Conserved hypothetical protein	√	+	-	-	+	-	-	-	-	-	-	35.2377	6.01	91.32	-0.285	1	14	318	-0.5
6. Proteins predicted by three groups of predictors (6)																					
1	PM0527	OM efflux protein TolC	√	+	+	+	+	+	+	-	+	+	+	50.6821	9.14	92.35	-0.317	2	21	455	10
2	PM0576	HemR	√	+	+	+	+	+	+	+	+	-	+	84.9103	9.24	72.4	-0.631	0	27	742	24.5
3	PM0646	LppC	√	-	-	+	+	+	+	-	+	+	+	63.3247	6.21	94.57	-0.217	0	26	571	0.5
4	PM0778	HexD	√	+	+	-	+	+	+	-	+	+	+	43.0036	9.45	102.7	-0.077	2	19	393	12
5	PM1016	Wza	√	+	+	-	+	+	+	-	+	+	+	42.2402	8.38	98.5	-0.091	1	19	387	6

Table 2-3. (Continued)

No	ID	Name	Presence in the porcine genome	Bioinformatic prediction of the OMPs										Physicochemical properties							
				Subcellular localization				TM β-barrel prediction				Lipoprotein									
				PA	PSORTb	SOSUI-GramN	CELLO	TMBETADISC-RBF	TMB-Hunt	BOMP	MCMBB	LIPO	LIPOP	MW (kDa)	PI	Aliphatic index	GRAVY score	No of TM helices	No of TM β-strands	Length (amino acids)	Charge
6	PM1323	Lipoprotein	√	+	-	-	+	+	+	-	+	+	+	29.7399	9.49	92.41	0.021	1	15	270	12
7. Proteins predicted by transmembrane β-barrel and lipoprotein predictors (2)																					
1	PM0674	Lipoprotein	√	-	-	-	-	+	+	-	+	+	+	25.2867	9.57	80.65	-0.36	0	11	230	14.5
2	PM1517	PlpE	-	-	-	-	-	+	+	-	+	+	+	37.4493	5.76	65.19	-0.798	0	13	335	-2

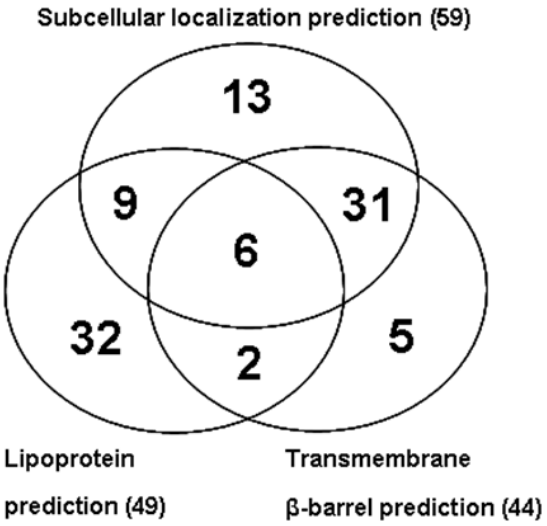


Figure 2-10. Comparative bioinformatic prediction of the 98 confidently predicted OMPs from the avian strain genome using three different groups of predictors (subcellular localization, transmembrane β-barrel protein and lipoprotein predictors). The predicted proteins in each group were determined as shown in **Figure 2-2**. The numbers represent shared or unique predicted proteins. The total number of proteins predicted by each of the three approaches is shown in parentheses.

Comparison of the DNA sequence identity of the confidently-predicted OMPs from the avian and porcine strain genomes indicated that the majority (64 proteins) of the predicted OMPs had sequence identities above 99% **(Figure 2-11)**. OMPs having DNA sequence identities less than 99% included HgbA (98%) and HgbB (haemoglobin and haemoglobin-haptoglobin receptors, 98%), OmpH_2 (a porin, 98%), NanH (sialidase, 98%), PM1717 (an autotransporter, 98%), LppA (98%), PilW (98%), TadD (97%), RcpA (96%), YccT (96%), FadL (95%), MltB (94%), OmpA (92%), NanB (89%), Hsf_2 (trimeric autotransporter, 87%), Hsf_1 (83%), PlpP (83%), LspB_2 (an autotransporter, 74%), PM0803 (TonB-dependent receptor, 63%), PM1543 (hypothetical protein, 63%), Opa (62%) and PlpE (56%).

2.3.5 Functions of the confidently predicted OMPs

The functions of the 98 confidently predicted OMPs in the avian strain genome are summarized in **Table 2-4**. These functions include outer membrane biogenesis and integrity (12 proteins), transport and receptor (25 proteins), adherence (7 proteins) and enzymatic activity (9 proteins). Forty-one proteins have unknown functions (although 17 are named) and 27 of these are lipoproteins. Interestingly, two or three copies of genes encoding certain proteins were predicted. For example, three *ompH* genes and two genes of *lspB, hsf, hgbB, plpE* and *hlpB* were predicted. Similar observations, including three *ompH* genes and two genes of *lspB, hsf, hgbA*, and *plpE*, were made for the porcine strain genome. Two proteins, HexD and Wza, were predicted from both genomes but they appear to have similar functions in capsular polysaccharide transport. Twelve TonB-dependent receptors including HemR (hemin receptor), PfhR and HasR (heme receptors), HmbR, HgbA and two HgbB (haemoglobin receptors), and PM0803, PM0745, PM1081, PM1282 and PM1428 were predicted in the avian strain genome; notably, most of these are involved in iron uptake. Similarly, 14 TonB-dependent receptors were identified in the porcine strain genome including HemR, PfhR and HasR, HmbR, two HgbAs, HgbB,

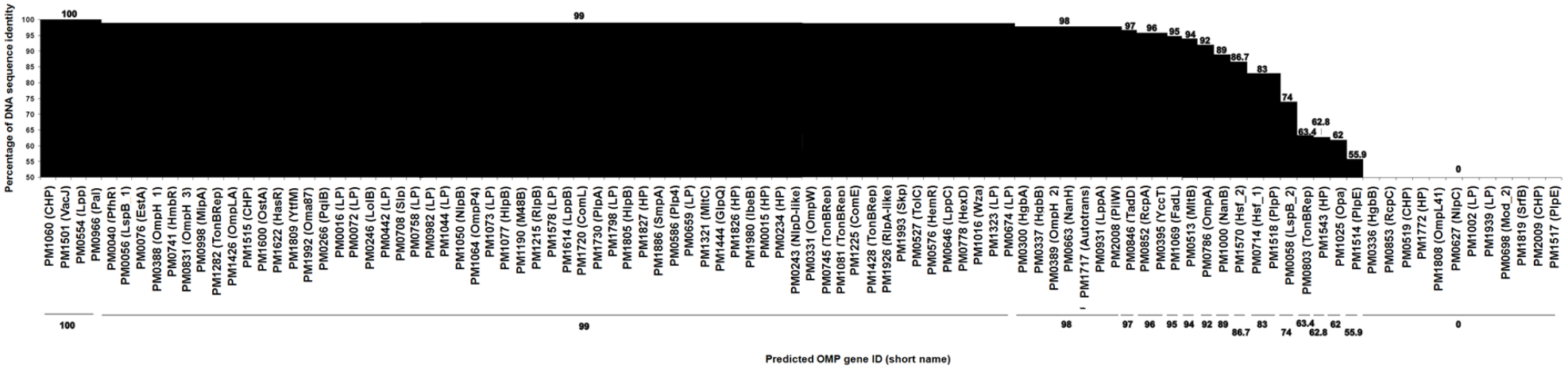


Figure 2-11. DNA sequences of the confidently predicted OMPs from the avian strain genome were compared to the confidently predicted proteins from the porcine strain genome using BLAST. The percentage of identity (y-axis) was plotted against the *P. multocida* avian strain gene IDs and short protein names in parentheses (x-axis). CHP, TonBRep, HP, LP and Autotrans are abbreviations for conserved hypothetical protein, TonB-dependent receptor, hypothetical protein, lipoprotein and autotransporter, respectively. Numbers above the graph indicate the percentage of identity and OMPs are grouped according to the same percentage of identity.

PM0803, PM1075, PM1081, PM1282, PM1428 and two uncharacterized porcine strain-specific proteins (PMpPor1882 and PMpPor2194).

2.3.6 Proteins that were left-out due to the criteria

This step was tested in the avian strain genome of *P. multocida*. Overall, the ten predictors identified 421 proteins out of 2015 proteins, accounting for 21% of the proteome. At the consensus prediction step, proteins predicted by the subcellular localization predictors and by the transmembrane β-barrel predictors were filtered out; however, no proteins were filtered out by the lipoprotein predictors (**Figure 2-5**). The criteria selected for the subcellular localization predictor group allowed 65 proteins to pass through, but filtered out 97 proteins. The criteria selected for the transmembrane β-barrel protein predictor allowed 47 proteins to pass through, but filtered out 282 proteins. Therefore, 60% of the predicted proteins were filtered out by the subcellular localization predictor groups and 86% were filtered out from the transmembrane β-barrel protein predictor group. Taken together, of the 421 predicted proteins, 339 proteins were filtered out by either the subcellular localization or transmembrane β-barrel protein predictors; **of these,** 40 proteins were filtered out by both predictor groups **(Figure 2-12)**. Further analysis of the 339 proteins by manual confirmation step revealed that 39 (12% of the filtered-out proteins) were putative OMPs and/or periplasmic proteins. However, 20 (6% of the filtered-out proteins) of these were predicted by the lipoprotein predictor group and were taken back into the confident list. Another six proteins (2% of the filtered-out proteins) passed the criteria of the subcellular localization predictor group but did not pass the criteria of the transmembrane β-barrel predictor group. Therefore, these were removed from the list of the filtered-out proteins. Thus, 13 proteins were filtered out (representing 4% of the filtered-out proteins) which might be true OMPs. Manual confirmation of these 13 proteins showed that seven were putative OMPs. These included HbpA/DppA (PM0592), NlpD (PM1507), RcpB (PM0851), MltA (PM0928), ComEA (PM1665), NlpI (PM1113)

Table 2-4. Functional classification of the 98 confidently predicted OMPs from the avian *P. multocida* genome

No	ID	Name	Prediction group[a]	Protein function
1. Outer membrane biogenesis and integrity				
1	PM0246	LolB	L	Chaperone & protein transport activity
2	PM0513	MltB	L	Cell wall catabolic process
3	PM1050	NlpB	L	Insertion of OMPs
4	PM1215	RlpB	L	LPS assembly
5	PM1614	LppB/NlpD	L	Cell wall catabolic process & proteolysis
6	PM1886	SmpA	L	Maintaining envelope integrity & β-OMP assembly
7	PM0786	OmpA	SB	Outer membrane integrity
8	PM0998	MipA/OmpV family protein	SB	MltA-interacting protein
9	PM1600	LptD/Imp/OstA	SB	LPS assembly/response to organic substance
10	PM1992	Oma87	SB	Outer membrane biogenesis & surface antigen
11	PM1321	MltC	SL	Cell wall catabolic process
12	PM0966	Pal/Omp P6	SL	Envelope integrity/link outer membrane to peptidoglycan
2. Transport and receptor				
1	PM0331	OmpW	S	Transport small hydrophobic molecules
2	PM0745	TonB-dependent receptor	S	Receptor & transporter activities
3	PM1081	TonB-dependent receptor	S	Receptor & transporter activities
4	PM1428	TonB-dependent receptor	S	Receptor & transporter activities
5	PM1720	ComL	L	DNA uptake/outer membrane biogenesis
6	PM1730	PlpA/MetQ	L	Amino acid transport
7	PM1069	FadL	SB	Transport hydrophobic compounds
8	PM1282	OM hemin receptor	SB	Haem receptor & transporter activities
9	PM0300	HgbA	SB	Hemoglobin receptor & iron transport
10	PM0336	HgbB	SB	Hemoglobin receptor & iron transport
11	PM0337	HgbB	SB	Hemoglobin receptor & iron transport

[a] Predictor groups; 'S' = subcellular localization predictors, 'B' = transmembrane β-barrel protein predictors, and 'L' = lipoprotein predictors.

Table 2-4. Continued

No	ID	Name	Prediction group	Protein function
12	PM1622	HasR	SB	Haem receptor & transporter activities
13	PM0741	HmbR	SB	Hemoglobin receptor & iron transport
14	PM0040	PfhR	SB	Hemoglobin receptor & iron transport
15	PM0803	TonB-dependent receptor	SB	Receptor & transporter activities
16	PM0388	OmpH_1	SB	Porin/ion transporter activity
17	PM0389	OmpH_2	SB	Porin/ion transporter activity
18	PM0831	OmpH_3	SB	Porin/ion transporter activity
19	PM0056	LspB_1	SB	Two-partner secretion/secretion of filamentous hemagglutinin
20	PM0058	LspB_2	SB	Two-partner secretion/secretion of filamentous hemagglutinin
21	PM1980	IbeB	SL	Lipid binding/transporter activity
22	PM0576	HemR	SBL	Haem receptor & transporter activities
23	PM0527	Outer membrane efflux TolC	SBL	Protein secretion/transporter activity
24	PM1016	Wza	SBL	Capsular polysaccharide transport
25	PM0778	HexD	SBL	Capsular polysaccharide transport
3. Adherence				
1	PM1225	ComE/PilQ	S	Pilus assembly/protein secretion
2	PM0852	RcpA	SB	Protein secretion/Flp pilus biogenesis
3	PM0853	RcpC	SB	Tight adherence & fibril production
4	PM0714	Hsf_1	SB	Adherence
5	PM1570	Hsf_2	SB	Adherence
6	PM1025	Opa	SB	Porin activity/adherence
7	PM0846	TadD	SL	Protein secretion/binding/assembly & transport of Flp pili
4. Enzymatic activity				
1	PM0243	NlpD-like protein	S	Metalloendopeptidase activity
2	PM0627	Lipoprotein NlpC/P60	L	Cell-wall peptidase
3	PM1190	Peptidase M48B family protein	L	Metalloendopeptidase activity/zinc ion binding
4	PM1064	Lipoprotein E/OmpP4	L	Acid phosphatase activity/utilization of NAD, NADP
5	PM1000	NanB	SB	Exo-alpha-sialidase/ produce free sialic acid as energy & carbon sources

Table 2-4. Continued

No	ID	Name	Prediction group	Protein function
6	PM0663	NanH	SB	Exo-alpha-sialidase/ produce free sialic acid as energy & carbon sources
7	PM0076	EstA	SB	Lipid metabolism/hydrolase activity, acts on ester bonds
8	PM1426	Phospholipase A/OmpLA	SB	Lipid metabolic process/maintain asymmetry of the OM
9	PM1444	GlpQ	SL	Glycerol metabolic process/lipid metabolic process
5. Others				
1	PM0698	Mod_2	S	DNA binding/N-methyltransferase activity
2	PM1819	Virulence factor SrfB	S	Unknown
3	PM1926	RlpA-like protein	S	Unknown
4	PM1993	Skp/Outer membrane p25	S	Unknown
5	PM2009	Conserved hypothetical protein	S	Unknown
6	PM0015	Hypothetical protein	S	Unknown
7	PM0234	Hypothetical protein	S	Unknown
8	PM1808	OmpL41/YtfN-like protein	B	Bacterial morphogenesis
9	PM0266	Mce/PqiB	B	Unknown
10	PM0395	YccT	B	Unknown
11	PM0519	Conserved hypothetical protein	B	Unknown
12	PM1772	Hypothetical protein	B	Unknown
13	PM0708	Slp	L	Starvation-inducible lipoprotein
14	PM1501	VacJ	L	Promoting spread of bacteria through tissues
15	PM1514	PlpE	L	Unknown
16	PM1518	PlpP	L	Unknown
17	PM1805	HlpB	L	Unknown
18	PM1077	HlpB	L	Unknown
19	PM2008	PilW/PilF	L	Unknown
20	PM0931	LppA	L	Unknown
21	PM0072	Lipoprotein	L	Unknown
22	PM0758	Lipoprotein	L	Unknown
23	PM0982	Lipoprotein	L	Unknown
24	PM1044	Lipoprotein	L	Unknown
25	PM1073	Lipoprotein	L	Unknown
26	PM1578	Lipoprotein	L	Unknown
27	PM1798	Lipoprotein	L	Unknown

Table 2-4. Continued

No	ID	Name	Prediction group	Protein function
28	PM0016	Lipoprotein	L	Unknown
29	PM0442	Lipoprotein	L	Unknown
30	PM1060	Conserved hypothetical protein	L	Unknown
31	PM1827	Hypothetical protein	L	Unknown
32	PM1002	Hypothetical protein	L	Unknown
33	PM1939	Hypothetical protein	L	Unknown
34	PM1717	Outer membrane autotransporter	SB	Unknown
35	PM1809	Omp85 family protein/YtfM	SB	Unknown
36	PM1515	Conserved hypothetical protein	SB	Unknown
37	PM1543	Hypothetical protein	SB	Unknown
38	PM0586	Plp4	SL	Unknown
39	PM0554	Lpp/Pcp	SL	Unknown
40	PM0659	Lipoprotein	SL	Unknown
41	PM1826	Hypothetical protein	SL	Unknown
42	PM1517	PlpE	BL	Unknown
43	PM0674	Lipoprotein	BL	Unknown
44	PM0646	LppC	SBL	Unknown
45	PM1323	Lipoprotein	SBL	Unknown

and a putative OMP (PM1623). The remainder were putative periplasmic proteins such as DctP (PM1651), ArtI (PM0124), LolA (PM0256). Thus, only seven (2%) of the left-out proteins were putative OMPs, while 332 (98%) were confidently removed by the selection criteria. The addition of these seven OMPs into the earlier list of 98 OMPs in the avian strain genome of *P. multocida* finally yielded 105 confidently predicted OMPs.

2.3.7 Physicochemical properties of putative OMPs

Analysis of physicochemical parameters **(Table 2-3)** highlighted the properties of the putative OMPs. The predicted proteins had molecular masses ranging from 7.1 to 276.2 kDa (52.4 $\pm$ 43 kDa average) and an average pI value of 8.1 $\pm$ 1.5. The average size of the predicted lipoproteins was smaller than that of the other proteins. Some proteins had very large sizes such as Hsf_1 (276 kDa) and the hypothetical lipoprotein PM0659 (214 kDa). The average GRAVY score (Kyte & Doolittle, 1982) was -0.35 $\pm$ 0.24 which indicated that the proteins were relatively hydrophilic compared to the predicted inner membrane proteins (data not shown). The predicted OMPs had more β-sheet strands (3-44 strands) than α-helices (0-3 helices).

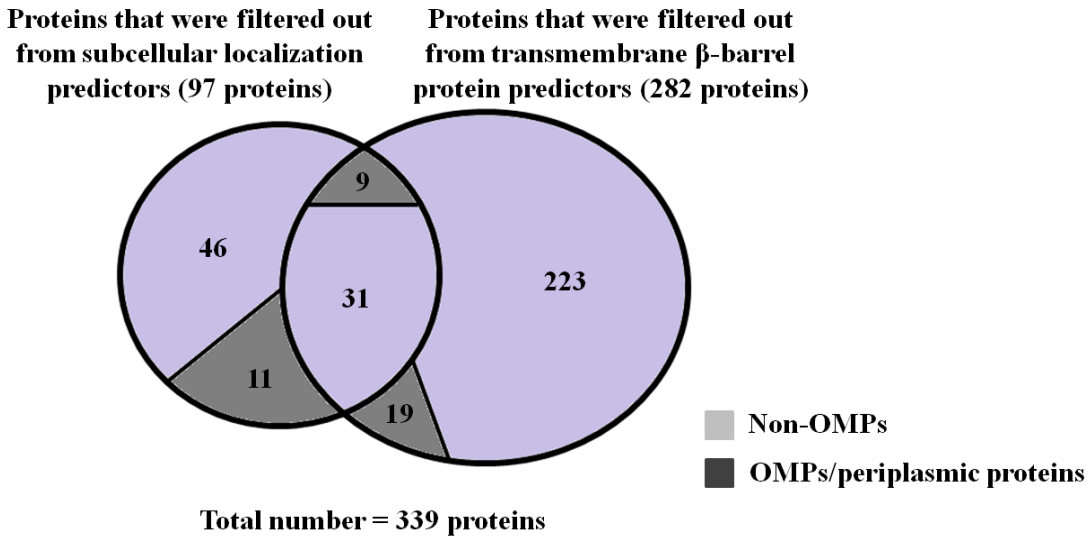


Figure 2-12. Comparison of proteins that were filtered out by the consensus criteria of the subcellular localization and the transmembrane β-barrel protein predictor groups. The aim of this additional analysis was to identify true OMPs that were lost after the consensus prediction. Area shaded in dark grey represents outer membrane or periplasmic proteins; area shaded in light grey represents non-OMPs.

2.4 Discussion

2.4.1 Different prediction methods

Each prediction method used in the present study **(Table 2-2)** is based on different algorithms and training datasets. The subcellular localization predictors aimed to determine all cellular components (secreted, outer membrane, periplasmic, inner membrane and cytoplasmic proteins) of the genome of *P. multocida*. PA analyzes keywords obtained from various databases using machine-learned classifiers and provides a user-friendly graphical explanation of each prediction (Szafron *et al.*, 2004). PSORTb combines multiple prediction components and each component performs a specific function including homology prediction, transmembrane prediction, a signal peptide prediction, and a specific motif prediction (Gardy *et al.*, 2003). SOSUI-GramN uses only the total sequence and physicochemical properties of the N- and C-terminal signal sequences for its prediction (Imai *et al.*, 2008). CELLO uses a supervised-learning method (support vector machines, SVMs) to detect specific amino acid compositions and motifs (Yu & Lin, 2004). Of 162 proteins predicted by the subcellular localization predictors from the avian strain genome, 15% were predicted by all four predictors, 25% by two or three predictors and 60% by a single predictor. Similarly, of 197 predicted proteins from the porcine strain genome, 11% of proteins were predicted by all four predictors, 26% by two or three predictors and 63% were predicted by a single predictor. Therefore, approximately 40% of the proteins predicted by the subcellular localization predictors were predicted by at least two predictors. Although PA and PSORTb have been widely used and reported as highly efficient predictors (Gardy & Brinkman, 2006), SOSUI-GramN and CELLO identified additional OMPs (e.g., RcpC, NanB, TadD, LppC and PM1515) which the first two predictors did not. The reason for this could be due to different algorithms used in SOSUI-GramN and CELLO. Overall, the use of multiple subcellular localization predictors increased both the prediction coverage and the confidence of prediction.

Conversely, the transmembrane β-barrel protein and lipoprotein predictors identified specific groups of OMPs. The four transmembrane β-barrel protein predictors discriminate between β-barrel proteins and non-β-barrel proteins. BOMP detects the C-terminal signal sequence and typical β-barrel pattern of the total amino acid sequence (Berven *et al.*, 2004). MCMBB uses a fast algorithm to determine alternating patterns of the transmembrane β-barrel proteins (Bagos *et al.*, 2004b). TMB-Hunt and TMBETADISC-RBF identify transmembrane β-barrel proteins based on amino acid composition profiles using different algorithms (Garrow *et al.*, 2005b; Ou *et al.*, 2008). MCMBB and TMB-Hunt predicted more proteins than BOMP and TMBETADISC-RBF **(Figure 2-6, Appendix Tables 2-1 and 2-2)**. The explanation for this could be due to differences in the algorithms, scoring schemes and performance levels. By using these four transmembrane β-barrel protein predictors, 30% and 33% of transmembrane proteins were predicted by at least two predictors from the avian and porcine strain genomes, respectively; the remaining transmembrane proteins were predicted by a single predictor. The use of multiple transmembrane β-barrel protein predictors again resulted in an increase in the confidence of prediction.

For the lipoprotein predictors, LIPO and LIPOP detect outer membrane lipoproteins by using their conserved lipo-box sequences. Together, LIPO and LIPOP predicted 65% of lipoproteins from the avian strain genome and 73% of lipoproteins from the porcine strain genome. These results indicate a high level of agreement between the two predictors and a high level of confidence.

Our findings confirm results obtained with *Escherichia coli* which showed that the use of multiple predictors increases the efficiency of subcellular localization prediction as well as specific-feature (β-barrel and lipid modified proteins) prediction when compared with the use of a single program (Díaz-Mejía *et al.*, 2009). Mirus and Schleiff compared different

transmembrane β-barrel protein predictors and showed that the combinatory approach improved the reliability of the prediction (Mirus & Schleiff, 2005). Moreover, we have also confirmed that the combined use of different predictors improves the coverage of predicted OMPs and our findings are consistent with previous work in other bacterial species (Berven *et al.*, 2006; Chung *et al.*, 2007; Huntley *et al.*, 2007). However, a higher number of predictors were used in the present study.

2.4.2 Filtration, integration and confirmation of the prediction results

In the present study, we used a combination of subcellular localization, transmembrane β-barrel protein and lipoprotein predictors, followed by consensus prediction with optimized criteria, integration and manual confirmation (data mining and sequence analyses) to predict OMPs in the available avian and porcine *P. multocida* genomes. The criteria stringency was optimized by maximizing precision, recall, specificity, accuracy and MCC. When we increased the stringency of the criteria **(Figure 2-9)**, such as from positive prediction by at least two predictors to three predictors, we observed a reduction of recall, meaning that most of the false-positives were removed but some true positives were possibly lost as well. Applying the consensus method and manual confirmation enhances the confidence and reliability of the predicted proteins (Díaz-Mejía *et al.*, 2009; Heinz *et al.*, 2009; Viratyosin *et al.*, 2008). Viratyosin *et al.* developed a computational framework incorporating consensus prediction of the subcellular localization predictors and homology information for subcellular localization prediction of the *Leptospira interrogans* genome and identified 63 putative OMPs (Viratyosin *et al.*, 2008). Similarly, Heinz *et al*. used multiple prediction phases, including screening of the inner membrane proteins, manual confirmation of the PSORTb database, and prediction of β-barrel, β-helix and lipoproteins, to identify the OMPs in *Chlamydiae* (Heinz *et al.*, 2009). Our study provides a simple framework which improves the confidence of prediction of the outer membrane proteome

of *P. multocida* compared to previous studies (Al-Hasani *et al.*, 2007; Hatfaludi *et al.*, 2010).

By using the consensus prediction followed by integration of the results for three predictor groups **(Figure 2-2)**, the number of predicted proteins decreased from 421 to 140 for the avian strain genome and from 439 to 147 for the porcine strain genome. Consensus prediction removed 332 proteins and only seven (2%) of these were confirmed as being putative OMPs. This indicated that the consensus prediction efficiently removed most of the false-positive proteins in exchange for a few putative OMPs. The manual confirmation step further reduced the numbers to 98 and 107 confidently-predicted putative OMPs for the avian and porcine *P. multocida* genomes, respectively. Combining seven filtered-out putative OMPs to the list of 98 OMPs of the avian strain genome resulted in 105 confidently-predicted putative OMPs. These values represent an average of 4.8% of the total proteome. The two predicted outer membrane proteomes were very similar, sharing 89 (83%) proteins. The majority (64) of these proteins had sequence identities above 99%, whereas 22 proteins had sequence identities in the range of 55.9-98%. Twelve proteins were present in either the avian or porcine genomes but not both. Of these, only three, namely, the Omp100 adhesin/invasin and two uncharacterized TonB-dependent receptors, were annotated as having putative function, in adherence and transport. The presence of these proteins in porcine isolates alone suggests a possible role in host adaptation.

Of the 98 confidently predicted putative OMPs from the avian strain genome, 48 proteins were predicted by at least two groups of predictors, while the remainder were identified by only one approach. We were able to classify the predicted OMPs into transmembrane β-barrel, lipidified transmembrane β-barrel, and lipidified proteins. The subcellular localization predictors predicted four potential β-barrel proteins that were filtered out by the β-barrel predictor group. The loss of these potential true OMPs in the prediction may

have occurred due to the stringent criteria used during the consensus prediction as increased stringency reduces the rate of false positives at the cost of an increased rate of false negatives. The manual confirmation of individual predicted proteins helped in the elimination of the false-positive proteins, such as some secreted and periplasmic proteins, and confidently confirmed that predicted proteins were targeted to the outer membrane. Moreover, it also assigned relevant functions to about 60% of the predicted proteins whose roles included outer membrane biogenesis and integrity, transport and receptor functions, adherence, and enzymatic activity. However, the remainder of the proteins, especially the lipoproteins, are hypothetical and require further characterization.

Eighty-six of the 105 (98+7) putative OMPs predicted from the avian strain genome in the present study were also identified in the previous study by Al-Hasani *et al*. (Al-Hasani *et al.*, 2007). These authors predicted 129 putative OMPs and secreted proteins from the avian *P. multocida* genome using only three predictors (PA, PSORTb and LIPOP). The additional 19 proteins **(Figure 2-13)** that we identified included seven proteins predicted by transmembrane β-barrel protein predictors (a pilus assembly protein RcpC, a sialidase NanB, Mce/PqiB, YccT, PM0519, PM1515, PM1772), three proteins predicted by lipoprotein predictors (PM1002, PM1798, PM1939), three proteins predicted by subcellular localization predictors (PM0015, PM0234, a RlpA-like protein PM1926), and one protein (PM1323) predicted by all these predictor groups. The remainder of five proteins (HbpA/DppA, RcpB, ComEA, NlpD and a hypothetical protein PM1623) were filtered out by the consensus prediction, but were added back to the list as shown in **section 2.3.6**. In contrast to the present study, Al-Hasani *et al*. did not apply consensus prediction to filter their predicted results and were interested in identifying both OMPs and secreted proteins (Al-Hasani *et al.*, 2007). Consequently, there was disagreement in the localization of 19 proteins between the three predictors (particularly between PA and PSORTb) and these proteins could not be concluded to be OMPs with certainty. Forty-

three proteins predicted by Al-Hasani *et al*. were not confidently predicted in the present study **(Figure 2-13)** (Al-Hasani *et al.*, 2007). Of these, 18 were not predicted and 25 were filtered out by consensus prediction or manual confirmation. Clearly, the use of a large number of predictors, together with consensus prediction, allowed us to identify a larger number of outer membrane-associated proteins with a greater degree of confidence.

Hatfaludi *et al.* reviewed the functions and classification of the OMPs of *P. multocida* and reported that 73 proteins were outer membrane-located based on previously published experimental research (Hatfaludi *et al.*, 2010). We have confidently predicted 48 of these proteins. Three of these proteins which included HbpA/DppA (PM0592), a competence-related DNA-binding and uptake protein ComEA, and Flp (Tad) operon protein RcpB were previously filtered out by the consensus prediction and re-added to the list according to **section 2.3.6**. Whereas 25 proteins were not predicted in the present study **(Figure 2-13).** One protein, TbpA, was not identified because of its absence from the avian and porcine strain genomes. The remaining 24 proteins were not included in our list of confidently predicted OMPs for a number of reasons. Six proteins were filtered out by consensus prediction (three proteins) or shown to be non-OMPs by manual confirmation (three proteins). The proteins that were filtered out by consensus prediction included a lipoprotein PM0979, an outer membrane-bounded sialic acid-binding protein NanP/YiaO, and an Flp (Tad) operon protein Flp1. The remaining 18 proteins were not identified as OMPs by any of the ten predictors in the present study. These included cytoplasmic proteins (3), inner membrane proteins (4), a periplasmic protein (1) and extracellular proteins (2). There are a number of explanations for the presence of these proteins in the list assembled by Hatfaludi *et al*. including contamination during outer membrane extraction and multiple subcellular localizations of certain proteins (Hatfaludi *et al.*, 2010). Significantly, of the 105 OMPs predicted from the avian strain genome in the present study, 57 OMPs **(Figure 2-13)** were not reported by Hatfaludi *et al.* (2010). These

included OmpH_3, Opa, Hsf_1 and _2, LolB, LppA, RlpB, PlpE, SmpA, Plp4, LppC, HexD and Wza. Clearly, these findings indicate that there is still a lack of experimental evidence relating to the structures and functions of the majority of the predicted outer membrane proteome.

Both Hatfaludi *et al*. and Al-Hasani *et al.* identified the same 44 proteins that were also predicted in the present study **(Figure 2-13)** (Al-Hasani *et al.*, 2007; Hatfaludi *et al.*, 2010). However, a further 42 proteins in our list were only predicted by Al-Hasani *et al.* whereas four protein were only reported by Hatfaludi *et al.* **(Figure 2-13)** (Al-Hasani *et al.*, 2007; Hatfaludi *et al.*, 2010). In the present study, we have predicted 15 proteins that were not described by Hatfaludi *et al*. or predicted by Al-Hasani *et al.* **(Figure 2-13)** (Al-Hasani *et al.*, 2007; Hatfaludi *et al.*, 2010). These include the Flp operon protein RcpC, the paraquat-inducible protein Mce/PqiB, YccT, NlpD, a RplA-like protein PM1986, and nine hypothetical proteins (PM1623, PM1515, PM0519, PM1772, PM1002, PM1798, PM1939, PM0015, PM0234 and PM1323). However, the functions of certain of these proteins have not been determined. Overall, the present study has improved the coverage of the predicted outer membrane proteome of *P. multocida* by 18% compared to that of Al-Hasani *et al*. (Al-Hasani *et al.*, 2007). Our simple prediction framework has allowed us to confidently predict and increase the coverage of the outer membrane sub-proteome of *P. multocida* by using currently available predictors and databases. Recently, Goudenege *et al*. created a subcellular localization database, CoBaltDB, for Bacteria and Archeae by incorporating 43 different predictors and 784 complete proteomes, but they did not give consensus localization of the predicted proteins and a decision for protein location has to be made by the users themselves (Goudenège *et al.*, 2010). By using this database, our prediction framework can also be applied to confidently predict subcellular localization in other bacterial species.

This study has designed a simple prediction framework that allows the prediction of putative OMPs from the available *P. multocida* genomes with a high level of confidence. The approach involves the use of multiple predictors divided into three groups, together with consensus prediction followed by integration and manual confirmation. This study has confidently identified 105 putative OMPs from the avian strain genome and 107 putative OMPs from the porcine strain genome of *P. multocida* with 83% overlap between the two genomes. The coverage of the outer membrane proteome of this bacterium has improved on previous research. The identification of previously unrecognized OMPs in strains of *P. multocida* from different host species will stimulate further studies into the molecular basis of the pathogenesis of this organism. This study not only provides a basis for furthering our understanding of the outer membrane proteome of *P. multocida* but can also be applied to other Gram-negative bacteria.

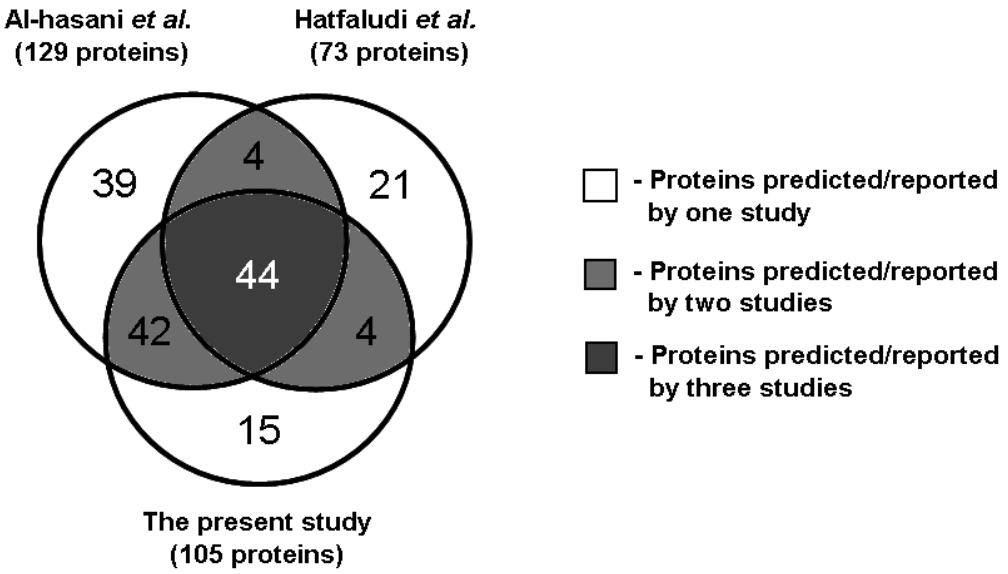


Figure 2-13. Comparison of the numbers of OMPs predicted in the present study with those predicted by Al-Hasani *et al.* (2007) and reported by Hatfaludi *et al.* .(2010). Indicated are the numbers of proteins predicted/reported by one, two or all three studies. The total number of proteins predicted/reported by each of the three studies is shown in parentheses.

Chapter 3: Comparative outer membrane proteomic analyses of *P. multocida* isolates from different host species

3.1 Introduction

Like all Gram-negative bacteria, the cell envelope of *P. multocida* consists of a symmetrical inner membrane and an asymmetrical outer membrane, separated by the periplasmic space and peptidoglycan layer (St Michael *et al.*, 2005). The outer membrane consists of an inner phospholipid layer and an outer LPS leaflet (Costerton *et al.*, 1974). The outer membrane harbours two classes of proteins, integral membrane proteins and lipoproteins, which together account for 2-3% of the total encoded proteins (Wimley, 2002). Integral membrane proteins typically have a β-barrel structure that traverses the membrane whereas lipoproteins are mostly anchored to the inner leaflet of the membrane (Costerton *et al.*, 1974; Schulz, 2002; Bos *et al.*, 2007). The outer membrane serves as a selective barrier that controls the passage of nutrients and waste products into and out of the cell and, crucially, provides the interface between pathogen and host. Thus, OMPs play important roles in the adaptation of bacteria to different environments and host niches (Ruiz *et al.*, 2006). Functions of *P. multocida* OMPs include biogenesis and integrity of the outer membrane (Omp87, OmpA, Lpp, Pal), nonspecific porin activity (OmpH_1 and _2), energy-dependent transport and binding activities (HgbA, HgbB, TbpA, HemR, HasR, PlpA/MetQ, TolC), adherence (ComE1, FhaB_1/LspB_1, FhaB_2/LspB_2, TadD, RcpA) and enzymatic activity (OmpLA, NanH, NanB, GlpQ) (Lin *et al.*, 2002; Kuhnert & Christensen, 2008; Hatfaludi *et al.*, 2010). Certain OMPs, such as PlpE, OmpH and FhaB2, are antigenic and used as protective immunogens for animals (Hatfaludi *et al.*, 2010). However, many of the OMPs of *P. multocida* remain uncharacterized and their functions are unknown.

3.1.1 Introduction to outer membrane proteomics

The majority of proteins destined for the outer membrane can be differentiated and predicted using bioinformatic approaches as described in **Chapter 2.** Bioinformatic predictors have been used to identify the numbers and functions of OMPs in *P. multocida* (Boyce *et al.*, 2006) and several other Gram-negative bacterial species (Berven *et al.*, 2006; Huntley *et al.*, 2007; Viratyosin *et al.*, 2008; Díaz-Mejía *et al.*, 2009). The outer membrane proteome has been characterized in a number of bacterial species using different combinations of proteomic technologies (Molloy *et al.*, 2000; Chung *et al.*, 2007; Cordwell *et al.*, 2008; Veith *et al.*, 2009). Two major approaches can be used to identify proteins - gel-based and non-gel-based proteomics (Poetsch & Wolters, 2008). Chung *et al.* (2007) characterized the outer membrane proteome of *Actinobacillus pleuropneumoniae* using a combination of the gel-based and gel-free proteomic approaches. These authors identified 50% of the predicted outer membrane proteome. In another study, Boyce *et al.* (2006) identified 24 OMPs from an avian strain of *P. multocida* using different the gel-based proteomic approaches. These authors used one-dimensional gel electrophoresis (1D-GE) followed by liquid chromatography mass spectrometry (LC-MS) and 2D-GE followed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS).

3.1.2 Extraction and separation of outer membrane proteins

Initially, appropriate strains are selected and cultured in suitable growth conditions with respect to osmolarity, nutrient composition, pH, temperature and aeration. Bacterial cells are harvested and washed with certain solutions such as buffered saline, Tris-HCl or EDTA (Hancock & Poxton, 1988). To extract the bacterial membrane, bacterial cells have first to be disrupted by either non-mechanical methods such as boiling in sodium dodecyl sulphate (SDS) or mechanical methods such as ultrasonic radiation or ultrasonication, explosive decompression, repeated high velocity compression and expansion by French Press,

Manton-Gaulin homogenizer, rapid agitation with small, rigid beads or bead beater, Braun homogenizer, Dynomill, Mini-Mill, Grinding, X-Press and Hughes Press cells (Hancock & Poxton, 1988). After the breakage of the bacterial cells, the cell envelopes are separated from the cytoplasmic proteins by centrifugation which pellets the cell envelopes and leaves the cytoplasmic proteins dissolved in the supernatant. Next, a variety of methods are used for separating the inner and outer membranes of the cell envelopes and enriching the proteins from a particular component. These include chemical solubilization and differential solubility using a series of solubilizing agents and centrifugation to remove insoluble material or differential centrifugation (density gradient centrifugation) to separate the two membranes into two layers due to their different densities (Cordwell, 2006). Okuyama *et al.* (1984) used density gradient centrifugation to separate the outer and inner membranes, and showed that the outer membrane density of *Vibrio* sp. was lower than that of the inner membrane due to the different lipid composition.

Extraction and purification of the OMPs from the outer membrane is a prerequisite for further analyses e.g. OMP profiling and proteomic identification. There are four main steps in the isolation, purification and identification of the OMPs from Gram-negative bacteria: outer membrane extraction, quantification of the OMPs, protein separation by SDS-PAGE and proteomic identification by mass spectrometry (Hancock & Poxton, 1988). Each of these is described separately below.

3.1.2.1 Outer membrane extraction methods

This step extracts and enriches the outer membrane fractions by separating them from the inner membrane and other cellular components. Different extraction methods can be used including detergent extractions and spheroplasting.

3.1.2.1.1 Detergent extraction

Detergents are a class of amphiphiles which are composed of a hydrophobic (lipophilic) tail and a hydrophilic head group (lipophobic). The hydrophilic group interacts with water molecules by hydrogen bonds, while the hydrophobic chains aggregate due to hydrophobic interactions and form spherical structures called micelles. The lowest concentration above which monomers gather to form micelles is defined as the critical micelle concentration (CMC) and the temperature at which the monomers reach the CMC is called the critical micellar temperature (CMT). Detergents solubilize membrane proteins by mimicking the lipid bilayer environment **(Figure 3-1)**. At low concentrations, detergents bind to the membrane by partitioning into the lipid bilayer. When the bilayers are saturated with detergents at high concentrations, the membranes disintegrate to form mixed micelles with the detergent molecules. Finally, mixed micelles containing lipids and detergents and detergent micelles containing proteins are formed. The performance of a detergent depends on various factors: detergent concentration, ionic strength, length of alkyl chain, pH, the presence of organic additives, purity and temperature. The hydrophile-lipophile balance (HLB) number is used to measure the hydrophilic character of the detergents: the larger the HLB, the more hydrophilic is the detergent. Non-denaturing detergents should have a HLB of between 12 and 20 (Bhairi, 2001).

Detergents can be broadly classified as ionic, non-ionic and zwitterionic **(Table 3-1)**. Ionic detergents have a head group with a net charge which can be either positively (cationic) or negative (anionic) charged. Examples include sodium N-lauryl sarcosine, sodium dodecyl sulfate, and bile acid salts (e.g., sodium salts of cholic acid and deoxycholic acid). The micelle size is influenced by the combined effect of the hydrophobic attraction of the side chains and the repulsive forces of the ionic groups, the concentration of counter ion and the increase in alkyl chain length. The ionic detergents

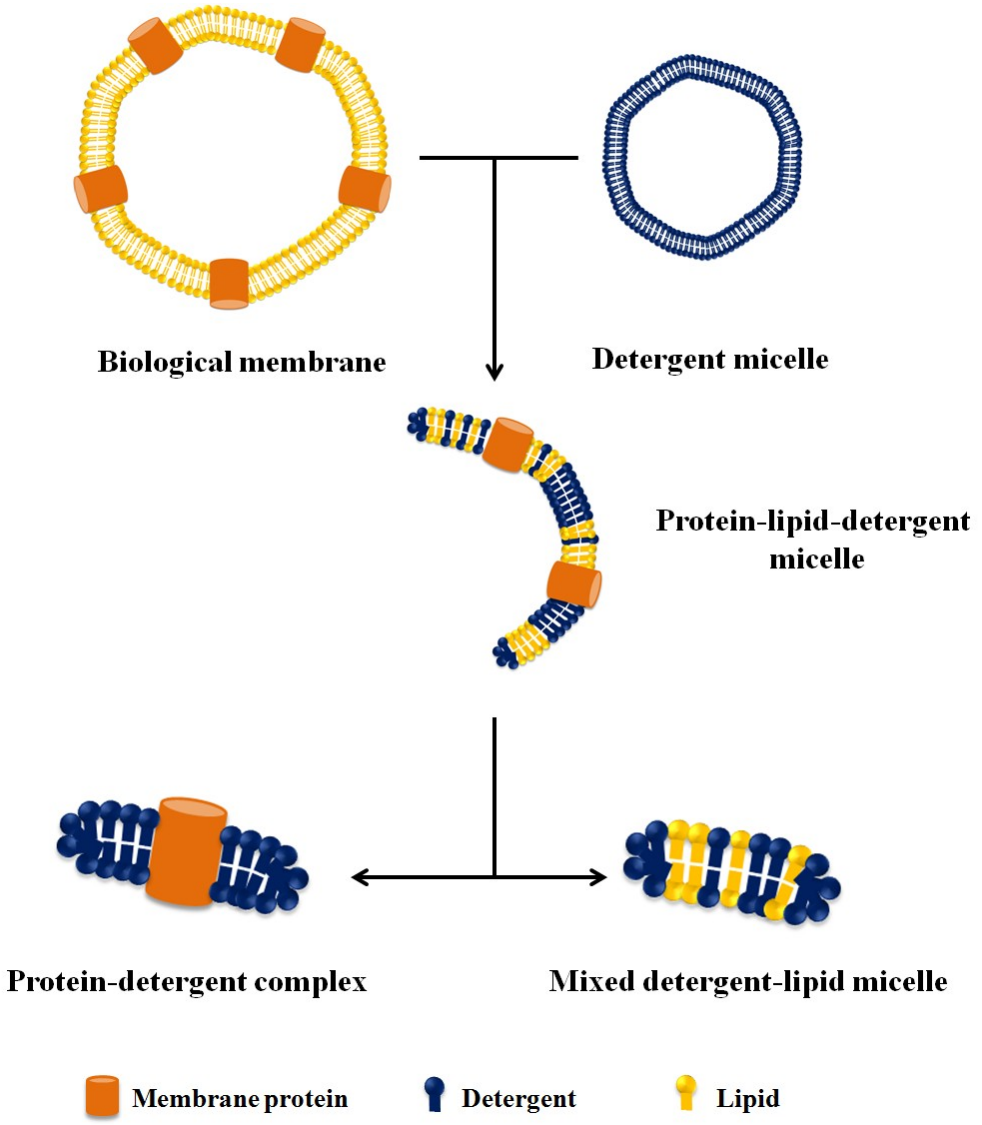


Figure 3-1. The use of detergents for membrane protein solubilisation (Bhairi, 2001). Detergents are amphiphilic molecules of which the hydrophilic group interacts with water molecules and the hydrophobic chains aggregate and form spherical structures called micelles. Detergents solubilize membrane proteins by mimicking the lipid bilayer environment. The mechanism of membrane protein extraction by detergent solubilisation is described in **section 3.1.2.1.1**.

are harsh and tend to be denaturing due to their efficiency of disrupting both inter- and intra-molecular protein-protein interactions (Bhairi, 2001).

Non-ionic detergents consist of uncharged, hydrophilic head groups. They are considered to be non-denaturing and are broadly used in the isolation of biologically active forms of membrane proteins. Examples include Triton X-100, NP-40, maltosides, glucosides, polyoxyethylene glycols. These detergents disrupt protein-lipid and lipid-lipid interactions better than protein-protein interaction. The non-ionic detergents become cloudy and undergo phase separation resulting in a detergent-rich layer and an aqueous layer at particular temperatures known as the cloud point. However, salts have minimal effect on the micelle size of the non-ionic detergents (Bhairi, 2001).

Zwitterionic detergents are electrically neutral and contain both positive and negative charges in their hydrophilic head groups. Examples are CHAPS/CHAPSO, Zwittergents and Fos-cholines. They are efficient in disrupting protein-protein interactions (Bhairi, 2001).

Treatment of cell envelopes with sodium lauryl sarcosinate or Sarkosyl will selectively solubilize the inner membrane leaving the outer membrane and peptidoglycan intact (Hancock & Poxton, 1988). Filip *et al.* (1973) proved this by the using isopycnic sucrose density gradient centrifugation to separate the inner membrane and outer membrane before solubilization with Sarkosyl, Triton X-100, SDS or Brij 58. From SDS-PAGE analysis, proteins solubilised by Sarkosyl in the supernatant were identical to those of the inner membrane. SDS completely solubilized both membranes, while Brij 58 slightly affected the inner membrane. Triton X-100 solubilized both membranes at concentrations ranging from 0.5-2%. The authors also found that Mg^{2+} prevented solubilization of the inner membrane by Sarkosyl.

Table 3-1. Classification and examples of detergents (Bhairi, 2001)

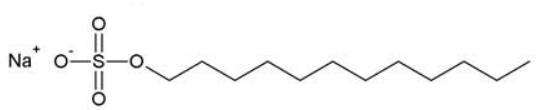
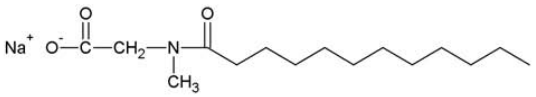
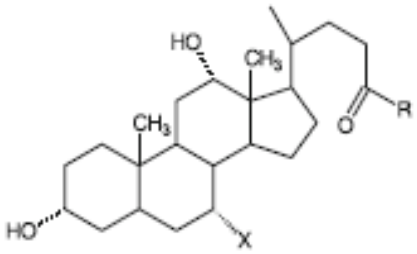
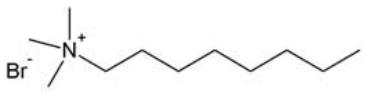
Detergents	Structures
1. Ionic detergents	
1.1 Anionic detergents	
1.1.1 Sodium dodecyl sulfate (SDS)	
1.1.2 N-lauryl sarcosine (Sarkosyl)	
1.1.3 Derivatives of bile acids	
(1) Sodium deoxycholate, x = H, R = O-Na^+ (2) Sodium taurodeoxycholate, x = H, R = $NHCH_2CH_2SO_3$-Na^+ (3) Sodium glycodeoxycholate, x = H, R = $NHCH_2CO_2$-Na^+ (4) Sodium cholate, x = OH, R = O-Na^+ (5) Sodium taurocholate, x = OH, R = $NHCH_2$ CH_2SO_3-Na^+ (6) Sodium glycocholate, x = OH, R = $NHCH_2CO_2$-Na^+	
1.1.4 Others (1) Chenodeoxycholic acid (2) 1-Octanesulfonic acid sodium salt (3) Sodium glycolithocholate (4) Sodium glycoursodeoxycholate (5) LPD-12	(6) Sodium taurochenodeoxycholate (7) Sodium taurodeoxycholate (8) Sodium tauroursodeoxycholate (9) Sodium ursodeoxycholate
1.2 Cationic detergents	
1.2.1 Octyltrimethylammonium bromide (OTAB)	
1.2.2 Others (1) Cetylpyridinium chloride monohydrate (CTAB) (2) Hexadecyltrimethylammonium bromide	

Table 3-1. (continued).

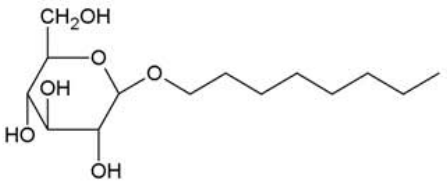
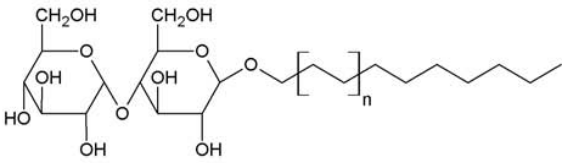
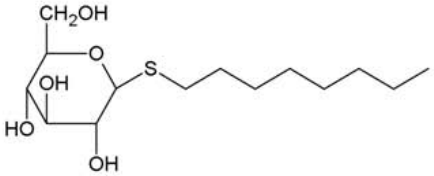
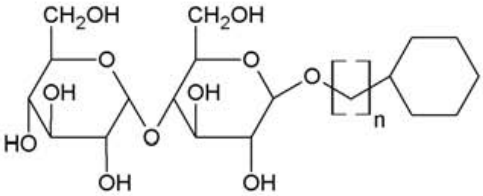
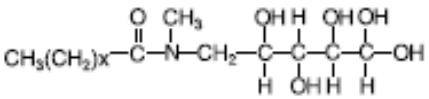
Detergents	Structures
2. Non-ionic detergents	
2.1 Alkyl glycosides	
	R-O-(CH2)x-CH3
	R = glucose
(1) N-nonyl-β-D-glucopyranoside, x = 8 (2) N-octyl-β-D-glucopyranoside, x = 7 (3) N-heptyl-β-D-glucopyranoside, x = 6 (4) N-hexyl-β-D-glucopyranoside, x = 5	
	R = maltose
(5) Dodecyl-β-D-maltoside, x = 11 (6) Decyl-β-D- maltoside, x = 9	
	R -S-(CH2)x-CH3
	R = glucose
(7) Octyl-β-D-thioglucopyranoside, x = 7	
(8) CYMAL-5, n = 5 (9) CYMAL-6, n = 6	
2.2 Glucamides	
(1) MEGA 10, x = 8 (2) MEGA 9, x = 7 (3) MEGA 8, x = 6	

Table 3-1. (continued).

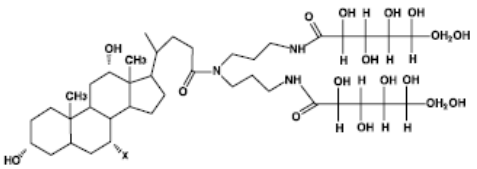
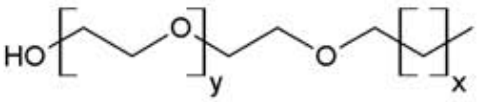
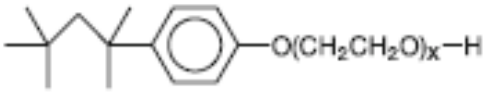
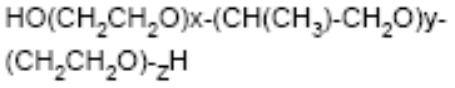
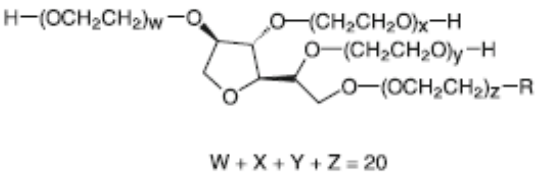
Detergents	Structures
(4) Deoxy big CHAP, x = H (5) Big CHAP, x = OH	
2.3 Polyoxyethylene detergents	
(1) BRIJ-35, x = 10, y = 22 (2) BRIJ-58, x = 15, y = 19 (3) Octyl polyoxyethylene (8-POE), x = 6, y = 0-10 (4) Pentaethylene glycol monooctyl ether (C8E5), x = 6, y = 4 (5) Hexaethylene glycol monooctyl ether (C8E6), x = 6, y = 5 (6) Octaethylene glycol monododecyl ether (C12E8), x = 10, y = 7 (7) Nonaethylene glycol monododecyl ether (C12E9), x = 10, y = 8	
(8) Triton X-100, NP-40, x = 9-10 (9) Triton X-114, x = 7-8	
(10) PLURONIC F-127, x = 98, y = 67, z = 98	
(11) TWEEN 20, R = C11H23CO2-(laurate) (12) TWEEN 80, R = C17H33CO2-(oleate)	

Table 3-1. (continued).

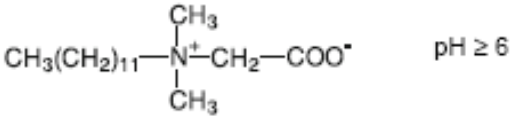
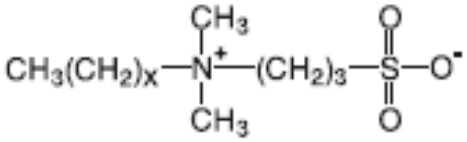
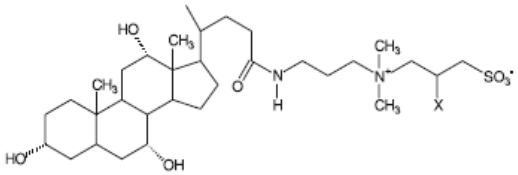
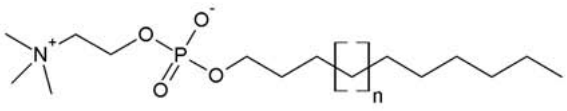
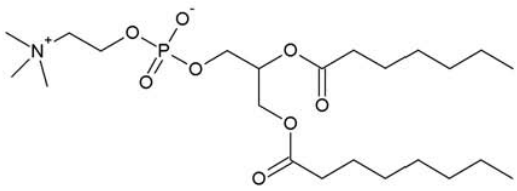
Detergents	Structures
2.4 Others	
(1) APO-10	(9) APO-12
(2) Cyclohexyl-n-hexyl-β-D-maltoside	(10) n-Decanoylsucrose
(3) Digitonin	(11) n-Dodecanoylsucrose
(4) ELUGENT Detergent	(12) GENAPOL C-100
(5) GENAPOL X-080, X-100	(13) HECAMEG
(6) n-Octanoylsucrose	(14) n-Octyl- β-D-glucopyranoside (OGP)
(7) n-Octyl- β-D-maltopyranoside	(15) n-Octyl- β-D-thioglucopyranoside
(8) Saponin	
3. Zwitterionic detergents	
(1) EMPIGEN BB (n-dodecyl-N,N-dimethylglycine)	 pH ≥ 6
(2) ZWITTERGENT 3-08, X = 7 (3) ZWITTERGENT 3-10, X = 9 (4) ZWITTERGENT 3-12, X = 11 (5) ZWITTERGENT 3-14, X = 13 (6) ZWITTERGENT 3-16, X = 15	
(7) CHAPS x = H, (8) CHAPSO x = OH,	
(9) N-decylphosphocholine (FOS-CHOLINE-10), n = 1 (10) N-dedecylphosphocholine (FOS-CHOLINE-12), n = 3	

Table 3-1. (continued).

Detergents	Structures
(11) Diheptanol phosphatidylcholine	
(12) Others	
(12.1) SB3-10	
(12.2) ASB-14	
(12.3) DDMAB	
(12.4) PMAL-B-100	
(12.5) SB3-12	
(12.6) ASB-16	
(12.7) DDMAU	

Sarkosyl has been widely used in the preparation of outer membranes from numerous bacterial species (Ravaoarinoro *et al.*, 1994; Marandi & Mittal, 1996; Brennan *et al.*, 1997; Peak *et al.*, 2000; Davies *et al.*, 2003a, 2003b, 2003d, 2004; Baik *et al.*, 2004; Rhomberg *et al.*, 2004; Hays *et al.*, 2005; Peng *et al.*, 2005; Xu *et al.*, 2005; Yagupsky & Slonim, 2005 Boyce *et al.*, 2006; Lee *et al.*, 2007). Sarkosyl-extracted OMPs of *P. multocida* have been studied by SDS-PAGE analysis (Marandi & Mittal, 1996; Davies *et al.*, 2003a, 2003b, 2003d, 2004; Boyce *et al.*, 2006). Brennan *et al.* (1997) isolated three OMPs (31 kD, 40 kD and 42 kD) of *Pasteurella haemolytica* A1 using Sarkosyl extraction and the authors concluded that this method yielded purer proteins than other methods including sucrose density gradient, isoelectrofocusing and chromatofocusing. Ravaoarinoro *et al.* (1994) compared Sarkosyl solubilization to isopycnic sucrose density gradient centrifugation for the isolation of OMPs of *P. aeruginosa*. They reported that Sarkosyl extraction yielded a higher OMP content and similar peptide pattern as sucrose gradient centrifugation.

Komatsuzawa *et al.* (2002) fractionated the outer membrane of *Actinobacillus actinomycetemcomitans* associated with periodontal disease by sucrose density gradient centrifugation and identified six Sarkosyl-insoluble OMPs (Omp100, Omp64, Omp39, Omp29, Omp18 and Omp16). Kim *et al.* (2006) applied Sarcosine extraction to the identification of closely related *Salmonella enterica* serotypes. They used Sarkosyl extraction to isolate the OMPs and dried them on a gold reflective slide. The samples were scanned using Fourier transform infrared (FTIR) spectroscopy. Spectra were analyzed using canonical variate analysis (CVA) and linear discriminant analysis (LDA). The authors stated that the use of the FTIR method combined with chemometrics provided better classification between bacterial isolates that had a high degree of similarity in the major OMP profiles.

However, Sarkosyl extraction may not yield a completely representative OMP profile (Anwar *et al.*, 1983; Stull *et al.*, 1985). Murphy and Loeb (1989) compared five different techniques of outer membrane isolation for *M. catarrhalis* and suggested that techniques based on selective detergent solubility of the outer and inner membranes were less efficient in isolating the outer membrane of *M. catarrhalis* compared to sucrose density gradient centrifugation. The method based on the collection of outer membrane vesicles, which included collection of vesicles from broth culture and EDTA-heat treated preparations, was more successful in isolating *M. catarrhalis* OMPs than Sarkosyl and TritonX-100 extractions. Their OMP profiles on polyacrylamide gels were similar to those obtained by sucrose gradient centrifugation.

In addition to Sarkosyl solubilisation, many other detergents or reagents have been used. Alkaline sodium carbonate was utilized in the preparation of OMPs from the soft-root phytopathogen *Dickeya dadantia* (Babujee *et al.*, 2007). Aivaliotis *et al.* (2004) isolated OMPs from the green sulphur bacterium *Chlorobium tepidum* by solubilization with Triton X-100. *Leptospira interrogans* OMPs were extracted using Triton X-114 and characterized by two-dimentional gel electrophoresis (2-DGE) and mass spectrometry (MALDI-TOF MS and tandem ESI-MS) (Cullen *et al.*, 2002). The OM of *Leptospira* is not tightly attached to the peptidoglycan and can be solubilized more easily than those of other Gram-negative bacteria (Cullen *et al.*, 2002).

Extraction of *A. pleuropneumoniae* OMPs using sucrose density gradient centrifugation followed by one of four different membrane wash treatments revealed that washing by sodium bromide paired with sodium carbonate (NBSC) yielded the most enriched OMPs (27 proteins) (Chung *et al.*, 2007). The authors suggested that sucrose density gradient

centrifugation followed by NBSC treatment was preferable to the sarkosyl-insoluble OM preparation.

3.1.2.1.2 Spheroplasting method

Bacterial spheroplasts are formed when the bacteria partially lose their cell envelope, including the outer membrane, resulting in sensitivity to low osmotic pressure (Voss, 1964). In Gram-negative bacteria, spheroplasts are prepared by digest the peptidoglycan with lysozyme in the presence of ethylenediaminetetraacetic acid (EDTA), tris-hydromethylaminomethane (Tris buffer) and sucrose. The use of an EDTA-saline wash can release the outer membrane fragments from the whole bacteria. The fragments form vesicles and these can be recovered after removal of the cells (Hancock & Poxton, 1988). This method has been used to prepare outer membranes in a number of bacterial species such as *Rhodobacter capsulatus* (Carmeli *et al.*, 1991), *Campylobacter jejuni* (Hobb *et al.*, 2009), *Yersinia pestis* (Pieper *et al.*, 2009) and *Francisella tularensis* (Huntley *et al.*, 2007).

Many Gram-negative bacteria produce outer membrane vesicles (~10 to 300 nm in diameter) naturally. These contain outer membrane and periplasmic components and can be produced in all stages of the growth and a variety of environments, particularly under stress conditions (Kuehn & Kesty, 2005; Ellis & Kuehn, 2010). These vesicles can function as a delivery system since they contain and deliver various virulence factors including protein adhesins, toxins, enzymes and LPS, to the host cells during infection. Lee *et al.* (2008) reviewed the identification of OMPs from the outer membrane vesicles of many Gram-negative bacterial species and found different OMPs such as OstA, OmpA, TolC, TonB-dependent receptors, porins OmpW and OmpF, bactericin-resistant factor OmpT, virulence factor IgA protease, nutrient uptake proteins LamB, BtuB and FadL,

murein hydrolases MltA, MltB and MltE, opacity protein and adhesin. The outer membrane vesicles may concentrate the virulence factors for targeted delivery and protect them from degradation. Moreover, these vesicles were also reported to contain DNA and antimicrobial agents such as autolysins which can degrade peptidoglycan of other bacteria (Mashburn-Warren *et al.*, 2008). Outer membrane vesicles can be obtained by centrifugation of the culture supernatant at low speed, filtration through a 0.45 µm filter and ultracentrifugation followed by sucrose density gradient centrifugation (Lee *et al.*, 2008). However, there is no evidence of naturally produced outer membrane vesicles in *P. multocida*.

3.1.2.2 Quantification of the extracted OMPs

Once the OMPs are extracted, protein quantification is the next important step which determines the protein yield and is essential for the separation step by electrophoresis. Several colorimetric protein assays have been widely used to quantify proteins and determine the protein concentration by reference to a standard curve of proteins with known concentrations. The modified Lowry assay (Markwell *et al.*, 1978) is a two-step assay in which the proteins are reacted with copper compounds in the first step and this protein-copper complex is then reduced. The Bradford assay is another method which involves protein-dye binding and detection of color change associated with the dye-bound protein (Bradford, 1976).

3.1.2.3 Protein separation methods

The isolated OMPs can be purified by gel permeation or chromatography techniques (Hancock & Poxton, 1988). The extracted OMPs are further purified and separated by either 1D or 2D SDS-PAGE resulting in the OMP profile which can be used to identify different strains and compare the outer membrane proteomes of Gram-negative bacteria.

3.1.2.3.1 One-dimensional SDS-PAGE

1D SDS-PAGE separates proteins on the basis of molecular mass (Westermeier, 2005). The protein samples are mixed with the anionic detergent SDS and heated at 100ºC before loading to disrupt the hydrogen bonds and unfold the proteins. The SDS molecules mask the charge of the proteins at a ratio of one gram of SDS per four grams of proteins. A reducing thiol agent such as 2-mercaptoethanol is also used to cleave disulfide bonds within the proteins. The 1D SDS-polyacrylamide gel is comprised of two regions: the stacking and resolving gels. The gels are submerged in an electrode buffer containing glycine which has no net charge and a low mobility, so it does not bind to the proteins. In the stacking gel, the SDS-bound proteins are concentrated and form stacks in the order of their mobility. The protein stack then moves constantly and slowly into the resolving gel towards the anode and the proteins are separated. The protein migration pattern depends on the gel structure, pH of the buffers (pH 6.8 for the stacking gel and pH 8.8 for the resolving gel) and the ionic strength of the buffer. Normally, the whole resolving gel has the same pore size, but the gradient gels can be prepared where the concentration of acrylamide varies throughout the gel; this improves protein banding over a wide molecular mass range.

3.1.2.3.2 Two-dimensional SDS-PAGE

2D SDS-PAGE separates proteins by two steps: isoelectric focusing in the first dimension and SDS-PAGE in the second dimension (Westermeier, 2005). The isoelectric focusing separate the proteins due to the isoelectric pH at which they are not charged. This first dimension is performed in gel strips which are next loaded onto the second dimension gels which further separate the proteins according to the molecular weight.

3.1.2.4 Identification of OMP using proteomic methods

The proteome is a complement of proteins expressed by an organism and the characterization of the complete set of proteins in a given organism is termed proteomics. Proteomic analysis generally involves several steps including protein extraction and separation, protein digestion, mass spectrometric analysis of peptides, data processing and protein identification. Mass spectrometry is an analytical technique that measures the mass-to-charge ratio of ions by generating a mass spectrum representing the relative masses of sample constituents (Westermeier & Naven, 2002; Veenstra & Yates, 2006).

3.1.2.4.1 Peptide sample preparation methods

There are two main analytical approaches used for preparation of peptides: gel-based and gel-free approaches (Westermeier & Naven, 2002; Veenstra & Yates, 2006).

3.1.2.4.1.1 Gel-based proteomics

The first approach uses gel electrophoresis to separate the proteome. 1D SDS-PAGE distinguishes proteins with respect to molecular masses while 2D-PAGE separates proteins by charge (isoelectric focusing) for the first dimension and molecular mass for the second. 1D SDS-PAGE can separate very hydrophobic membrane proteins better than 2D-PAGE although the resolving power is less (Weiner & Li, 2008). 2D-PAGE has limitations such as an inability to separate highly hydrophobic proteins and highly alkaline proteins (O'Connor & Hames, 2008). The protein spots can be visualized by a variety of methods including organic dye staining such as Coomassie Blue, silver nitrate staining, radioactive labelling, immunoblotting and fluorescent-based staining such as SYPRO Orange, SYPRO Red and SYPRO Ruby (Veenstra & Yates, 2006). The gel-based method also provides quantitative information on proteins so it is convenient to compare the proteomes of different samples. After separation, protein spots of interest are excised and in-gel digested by a proteolytic enzyme such as trypsin or a chemical reagent such as cyanogen bromide (CNBr). The resultant peptides are analyzed by a mass spectrometer (MS) and

peptide mass spectrums are matched with predicted peptides in genome databases to identify the proteins. However, gel bands from 1D SDS-PAGE may contain mixtures of proteins due to limited separation. This problem can be solved by the use of MS/MS or liquid chromatography LC-MS/MS (see **section 3.1.2.4.2**) (Weiner & Li, 2008). 1D SDS-PAGE LC-MS/MS has been used for the separation and identification of bacterial surface proteins (Cordwell, 2006; Bridges *et al.*, 2008; Weiner & Li, 2008).

3.1.2.4.1.2 Gel-free proteomics

The gel-free method digests the entire protein mixture and the resulting peptides are separated and analyzed by LC MS/MS. Selection of an appropriate solution to solubilise the proteins is important and the solvent should not totally denature the enzyme during digestion. SDS can be used to dissolve membrane proteins but high concentration of SDS will denature trypsin and is difficult to remove. Organic acids such as trifluoroacetic acid (TFA) and organic solvents such as methanol are useful in the solubilisation of membrane proteins. The gel-free approaches can identify more rare and less soluble proteins than the gel-based approach. Nevertheless, the limitation of the gel-free method is that it does not directly provide quantitative information of relative protein abundance. Differential tagging approaches are optional for quantitative studies (Burchmore, 2006; Weiner & Li, 2008).

Shotgun proteomics or multidimensional protein identification technology (MudPIT) uses 2D-LC which comprises strong cation exchange (SCX) in the first dimension and reverse-phase chromatography in the second followed by MS/MS (Cordwell, 2006). This method is less biased against highly hydrophobic proteins and can overcome some limitations of 2-DE. The use of microwave-assisted acid hydrolysis (MAAH) with TFA can degrade proteins into peptides for MS and does not require the use of solvent to dissolve proteins and enzyme digestion. However, a combination of several methods (Wang *et al.*, 2007)

including methanol-assisted trypsin digestion, SDS-assisted trypsin digestion and MAAH followed by LC-ESI MS/MS may result in better proteome coverage.

3.1.2.4.2 Protein identification using mass spectrometry

Mass spectrometer analyses sample molecules in the form of ionized gas. The mass spectrometer consists of three major components: 1) ionization sources where the samples are ionized and desorbed into a gas phase, 2) mass analyzers that guide the gas-phase ions, and 3) the detector. Two methods are typically used to ionize peptide samples. Matrix-assisted laser desorption ionization (MALDI) generates ions by irradiating the sample co-crystallized in organic matrix compound with a pulsed laser beam whilst electronspray ionization (ESI) uses a high voltage to ionize the sample dissolved in solution and the sample is desorbed before entering the analyzer region. Various types of mass analyzers have been developed including the ion-trap mass spectrometer, time-of-flight mass spectrometer (TOF), triple quadrupole mass spectrometer, quadrupole time-of-flight mass spectrometer (QqTOF) and Fourier transform ion cyclotron resonance mass spectrometer (FTICR). Tandem mass spectrometry (MS/MS) consists of more than one analyzer and can be used in structural and sequencing studies (Westermeier & Naven, 2002; Veenstra & Yates, 2006). With the MS/MS, the peptides are selected with the first mass analyzer and fragmented before the second analysis by the next mass analyzer (Westermeier, 2005).

3.1.2.4.2.1 MALDI TOF-TOF MS

This mass spectrometry method identifies proteins by using peptide mass finger printing (Westermeier, 2005). The peptide samples are mixed with low molecular weight compounds, called the matrix such as α-cyano-4-hydroxy cinnamic acid, which absorbs maximum energy at the wavelength of the laser. The mixtures are spotted and dried onto metal slides. Once placed into the mass spectrometer, the laser beam is fired into the peptide-matrix samples which absorb the energy and move away from the metal plate.

Then, the matrix molecules transfer their charges to the peptides and these peptides travel along the vacuum tubes at different velocity due to the mass. The flight times of the peptide ions to the detector are used to calculate the mass-to-charge ratio.

3.1.2.4.2.2 LC-ESI-Q/TOF MS

This mass spectrometry method is normally coupled with liquid chromatography (LC) system which separates the peptide mixture before applying it to the mass spectrometer. The samples are sprayed through a metal capillary as fine highly charged droplets. Next, the droplet size is reduced until the peptide ions can leave the droplet and enter the mass analyzer.

3.1.3 Objectives

In previous studies, the OMP profiles of 466 *P. multocida* isolates recovered from diseased cattle (153), sheep (55), pigs (158) and poultry (100) in England and Wales were characterized (Davies *et al.*, 2003a, 2003b, 2003c, 2004). In each case, a restricted number of clones were responsible for the majority of infection in each species (Davies *et al.*, 2004). The dominant clonal groups associated with each host possessed unique OMP profiles based on the molecular mass heterogeneity of the two major OMPs, OmpA and OmpH, together with variation in the expression and size of minor proteins. These observations suggested that differences in the outer membrane proteomes of strains from different hosts correlate with differences in host specificity and disease pathogenesis (Davies *et al.*, 2004). However, the identity of the majority of the OMPs present in these profiles is currently unknown. Clearly, the identification of these proteins, together with improved knowledge about their variation among isolates, are key to a better understanding of the molecular interactions that occur between *P. multocida* and its various animal hosts which ultimately lead to disease. This chapter aimed to characterize and compare the outer membrane sub-proteomes of eight isolates representative of major

clonal groups of *P. multocida* associated with different disease syndromes in cattle, sheep, pigs and poultry. Complementary proteomic methods, including gel-based and gel-free techniques, were used to identify candidate OMPs in the eight isolates. These proteins were also compared to the predicted outer membrane proteome described in **Chapter 2**.

3.2 Materials and methods

3.2.1 Bacterial isolates and growth conditions

Eight representative isolates of *P. multocida* recovered from infected cattle (two), sheep (two), pigs (two) and poultry (two) were investigated in this study (**Table 3-2**). The isolates were carefully selected to represent major clonal groups associated with disease in each host (Davies *et al.*, 2003b, 2004, 2003d, a). The avian isolates were recovered from cases of septicaemia (fowl cholera); the bovine and ovine isolates were recovered from cases of pneumonia; and the porcine isolates were recovered from cases of pneumonia and atrophic rhinitis. The evolutionary relationships of the isolates based on multilocus sequence typing (MLST) data (http://pubmlst.org/pmultocida_multihost/) were also taken into consideration (**Figure 1-3**). The isolates were stored at -80ºC in 50% (v/v) glycerol in brain heart infusion broth (BHIB; Oxoid). From -80ºC stock cultures, bacteria were streaked onto blood agar (brain heart infusion agar containing 5% (v/v) defibrinated sheep's blood) and incubated overnight at 37ºC. For preparation of outer membranes, the isolates were cultured overnight in 10 ml volumes of BHIB at 37ºC on an orbital shaker at 120 rpm. Eight hundred microlitre volumes of the overnight growth were inoculated into pre-warmed 400 ml volumes of BHIB in 2-litre non-dimpled Erlenmeyer flasks and incubated until the OD reached 0.8-0.9 (late-log phase) at 37ºC on an orbital shaker at 120 rpm.

Table 3-2. Properties of eight representative isolates of *P. multocida*.

Isolate	Animal host	Capsular serotype	OMP type[a]	Disease
PM144	Avian	A	1.1	Septicaemia
PM246	Avian	F	2.2	Septicaemia
PM564	Bovine	A	2.1	Pneumonia
PM632	Bovine	A	4.1	Pneumonia
PM684	Porcine	A	6.1	Atrophic rhinitis
PM734	Porcine	A	1.1	Pneumonia
PM966	Ovine	A	1.1	Pneumonia
PM982	Ovine	D	3.1	Pneumonia

[a)] OMP-typing schemes have been described separately for avian, bovine, porcine and ovine isolates (Davies *et al.*, 2003a, 2003b, 2003c, 2004)

3.2.2 Preparation of OMPs

3.2.2.1 Standard method

Outer membrane proteins were prepared by Sarkosyl extraction according to the method previously described by Davies *et al.* (Davies *et al.*, 2003a, 2003b, 2003d, 2004). Briefly, bacterial growth was halted by chilling in iced-water for 5 min and the bacterial cells were harvested by centrifugation at 13,000 x *g* for 20 min at 4ºC. The pelleted bacteria were resuspended in 50 ml of 20 mM Tris/HCl (pH 7.2) and centrifuged at 12,000 x *g* for 30 min at 4ºC. The sedimented cells were resuspended in approximately 7 ml of ice-cold 20 mM Tris/HCl (pH 7.2) and sonicated in iced-water for 5 min with a Soniprep sonicator (12 microns amplitude). The sonicated samples were adjusted to 10 ml and centrifuged at 12,000 x *g* for 30 min at 4ºC to remove unbroken cells. The supernatants were centrifuged at 84,000 x *g* for 1 h at 4ºC in a Sorvall ultracentrifuge to pellet the cell envelopes. The gelatinous pellets were rigorously resuspended in a final volume of 10 ml 0.5% sodium N-lauroylsarcosine (Sarkosyl; Sigma) for 20 min at room temperature to completely solubilize the cytoplasmic membranes and centrifuged at 84,000 x *g* for 1 h at 4ºC to pellet the outer membranes. The gelatinous outer membranes were resuspended in 20 mM Tris/HCl (pH 7.2) and centrifuged at 84,000 x *g* for 1 h at 4ºC. The final pellets were resuspended in approximately 1 ml of 20 mM Tris/HCl (pH 7.2). Fifty microlitre aliquots of the outer membrane preparation were transferred to Eppendorf tubes and protein concentrations were determined by the modified Lowry procedure (Markwell *et al.*, 1978). One hundred microlitre aliquots of the outer membrane samples were adjusted to 2 mg/ml. The outer membrane fractions were stored at -80ºC.

3.2.2.2 Preparation of OMPs using different detergents

First, optimal concentrations of Sarkosyl were determined by using the bovine isolate PM632 and the porcine isolate PM684. The outer membrane-enriched fractions were

prepared by the same method described in the **section 3.2.2.1**, but the concentrations of Sarkosyl used were 0.1%, 0.5%, 1% and 2%. Second, the ability of different detergents in extraction of the outer membrane fraction was tested on the bovine isolate PM632. The outer membrane fractions were prepared as described in **section 3.2.2.1**, but different detergents were used instead of 0.5% Sarkosyl. These detergents included 2% Triton X-100, 1% Sarkosyl + 7mM EDTA, 1% Triton X-114, 0.5% CHAPS, 1% Octylglucoside, 1% Deoxycholate. Carbonate extraction was also used.

3.2.2.3 Preparation of OMPs using spheroplasting method

Methods for producing spheroplasts were optimized from Coquet *et al*. (2005) and Khemiri *et al.* (2008a). Two isolates were selected: bovine isolate PM632 and porcine isolate PM684. Bacterial cells were harvested by centrifugation at 14,000 x g for 30 min and washed with 20% (w/v) sucrose. Cells were kept on ice. Approximately 1.5 g wet weight of cells were resuspended in a digestion solution that contained 2 M sucrose, 0.1 M Tris-HCl (pH 7.8), 1% EDTA (pH 7.0) and 0.5% lysozyme. The mixture was incubated for 1 h at 37°C in the presence of DNAse. Spheroplasts were removed by centrifugation at 10,000 x g for 20 min at 4°C. The outer membrane-containing fractions were collected by ultracentrifugation at 126,000 x g for 1h at 4°C. Proteins within supernatants were precipitated by cold-acetone precipitation. Briefly, 800 µl of cold acetone were added to 200 µl volume of the sample. The samples were vortexed and incubated for 16 h at -20 ºC. After incubation, the samples were centrifuged at 13,000 x g for 5 min. The sample pellets were washed in 80% acetone and the pellets were stored at -80ºC.

3.2.2.4 Preparation of cell envelope proteins

The eight isolates were grown and harvested using the same method described in **section 3.2.2.1**. Briefly, bacterial growth was halted by chilling in iced-water for 5 min and the bacterial cells were harvested by centrifugation at 13,000 x *g* for 20 min at 4ºC. The

pelleted bacteria were resuspended in 50 ml of 20 mM Tris/HCl (pH 7.2) and centrifuged at 12,000 x *g* for 30 min at 4ºC. The sedimented cells were resuspended in approximately 7 ml of ice-cold 20 mM Tris/HCl (pH 7.2) and sonicated in iced-water for 5 min with a Soniprep sonicator (12 microns amplitude). The sonicated samples were adjusted to 10 ml and centrifuged at 12,000 x *g* for 30 min at 4ºC to remove unbroken cells. The supernatants were centrifuged at 84,000 x *g* for 1 h at 4ºC in a Sorvall ultracentrifuge to pellet the cell envelopes. The gelatinous cell envelopes were resuspended in 20 mM Tris/HCl (pH 7.2) and centrifuged at 84,000 x *g* for 1 h at 4ºC. The final pellets were resuspended in approximately 1 ml of 20 mM Tris/HCl (pH 7.2) and stored at -80ºC.

3.2.3 Protein separation

All protein samples from the **section 3.2.2** were separated using 1D SDS-PAGE. However, 2D SDS-PAGE was used to examine the OMPs of bovine isolate PM632.

3.2.3.1 1D SDS-PAGE

Twenty micrograms of OMPs were separated by 1-D SDS-PAGE in 12% or 15% (w/v) linear gels or in 8-20% (w/v) gradient gels using the SDS discontinuous system of Laemmli (Laemmli, 1970) and the Hoefer SE600 electrophoresis apparatus as previously described (Davies *et al.*, 2003a, 2003b, 2003d, 2004). The amount of sample loading was increased in some gels (40-50 µg) to visualize proteins of low abundance. The gels were run until the dye front had reached the bottom of the resolving gel or for longer periods to obtain better separation of high-molecular-mass proteins. Proteins were visualized by staining with Coomassie brilliant blue R250.

3.2.3.2 2D SDS-PAGE

The bovine isolate PM632 was selected to compare three different protocols for 2D SDS-PAGE. Briefly, three samples of Sarkosyl-extracted outer membrane samples from the

section 3.2.2.1 were resuspended in 20 mM Tris-HCl (pH 7.2) containing 10 mM EDTA and stirred continuously for 16 h at 4ºC. The second sample was boiled for 5 min prior to the EDTA treatment, whereas the first and third samples were not boiled. After that, all three samples were precipitated by cold acetone for 16 h at 4 ºC. The precipitated samples were pelleted and resuspended in appropriate rehydration buffer as follows. Next, the first sample was resuspended in 470 µl of rehydration buffer containing 6 M urea, 2 M thiourea, 4% v/v CHAPS, 0.02% w/v DTT, 0.02% w/v bromophenol blue and 0.5% v/v IPG buffer. The second sample was resuspended in 470 µl of rehydration buffer containing 7 M urea, 2 M thiourea, 1% v/v amidosulfobetaine-14 (ASB-14), 2 mM tributylphosphine, 0.02% w/v bromophenol blue and 1% v/v IPG buffer. The third sample was resuspended in the 470 µl of rehydration buffer containing 7 M urea, 2 M thiourea, 1% v/v ASB-14, 4% CHAPS, 2 mM tributylphosphine, 0.02% w/v bromophenol blue and 1% v/v IPG buffer. These three samples were loaded into the IPG strip holder together with dehydrated pH 3-10 nonlinear IPG strips (Amersham Bioscience, UK) and covered with Drystrip cover fluid. The samples were isoelectric focussed using the IPGPhor II machine (Amersham Bioscience, UK). The strips were rehydrated at 20 ºC for 10-15 h and isoelectric focusing was carried out at 50 mA per strip; applying 300 V for 300 V·h, 1,000 V for 1,000 V·h, a linear voltage increase to 8,000 V (3975 V·h) and 8000 V for 70,000 V·h. When the isoelectric focusing was completed, the strips were equilibrated for 15 min in 10 ml of equilibration buffer containing 75 mM Tris-HCl (pH 8.8), 6 M urea, 30% w/v glycerol, 2% w/v SDS, 0.01 w/v bromophenol blue and 100 mg of dithiothreitol (DTT). The strips were further equilibrated for 15 min in 10 ml of the equilibration buffer containing 250 mg of iodoacetamide (IAA) instead of DTT. Equilibrated gel strips were placed on top of a vertical slab gel and held in place by the addition of molten agarose. The gels were then loaded to in a DALT 12 gel tank (Amersham Bioscience, UK) filled with electrophoresis buffere containing 25 mM Tris-HCl (pH 8.3), 193 mM glycine and

0.2% w/v SDS. The gels were run at 2 W per gel at 15 ºC until the dye front reached the bottom of the gel. Once finished, the gels were fixed in a solution containing 40% v/v ethanol and 10% v/v acetic acid for 1 h, and then stained with colloidal Coomassie.

3.2.4 Proteomic analyses

3.2.4.1 Preparation of peptides

3.2.4.1.1 Gel-based proteomic analysis

For mass spectrometric analysis, peptides were prepared by in-gel trypsin digestion as previously escribed by Szöor *et al.* (Szöor *et al.*, 2010). Gel bands were excised and each gel piece was placed in a separate well of a 96-well plate. Using the Ettan Spot Handling Workstation (Amersham Biosciences), the gel slices were washed in 500 µl of 100 mM ammonium bicarbonate followed by 50% acetonitrile/ 100 mM ammonium bicarbonate. The gel slices were incubated with 10 µl of 45 mM dithiothreitol (DTT) and 150 µl of 100 mM ammonium carbonate at 60ºC for 30 min to reduce disulfide linkages. Next, 10 µl volumes of 100 mM iodoacetamide (IAA) were added and incubated at room temperature in the dark for 30 min to aminocarboxymethylate cysteine residues. The gel slices were subsequently washed once with 500 µl volumes of 50% acetonitrile/ 100 mM ammonium bicarbonate, shrunk in 50 µl volumes of acetonitrile, and dried in a vacuum centrifuge. The gel slices were rehydrated with approximately 20 µl of trypsin solution (0.2 µg/ml sequencing grade modified porcine trypsin in 25 mM ammonium bicarbonate) and incubated overnight at 37ºC. Next, equal volumes of acetronitrile were added to the digests and incubated for 20 min, then all liquid was transferred to a 96-well microtitre plate. Additional peptides were extracted from the gel pieces by two further incubations in sufficient volumes of 1% formic acid (20 min) to cover the gel pieces followed by two further extractions in acetronitrile (10 min). The peptide extracts from each sample were combined in the same well of a microtitre plate and dried in a vacuum centrifuge. Dried

peptide samples were stored at -20ºC and were resuspended in 4 µl of 50% (v/v) acetronitrile and 0.1% (v/v) trifluoroacetic acid (TFA) prior to MALDI-TOF-TOF MS analysis (**section 3.2.4.2.1**).

3.2.4.1.2 Gel-free proteomic analysis

Two methods of preparation were employed: methanol-aided trypsin digestion modified from Bridges *et al.* (Bridges *et al.*, 2008) and phase transfer surfactant-aided trypsin digestion (Masuda *et al.*, 2008). For the methanol-aided trypsin digestion, 4 mg/ml of the outer membrane samples were resuspended in 44 µl of 50 mM ammonium bicarbonate then placed in a sonicator bath for 20 min, with mixing at five minute intervals. The samples were incubated for 20 min at 60ºC and then placed on ice for 3 min. Sixty microlitres of methanol were added and the samples were incubated for 5 min in the sonicator bath and mixed. Sixteen microlitres of trypsin solution (20 µg/ml in 25 mM ammonium bicarbonate) and 60 µl of methanol were added (to give a final concentration of 60% [v/v] methanol) and the samples were mixed. After incubating for 12-16 h at 37ºC, the digested samples were dried down to a 5 µl volume in a vacuum centrifuge and stored at -20ºC until LC-ESI-Q/TOF MS analysis. The samples were resolubilized in 20 µl of 2% acetronitrile and 0.1% formic acid prior to LC-ESI-Q/TOF MS analysis (**section 3.2.4.2.2**).

For the phase transfer surfactant-aided trypsin digestion, 25 µl of the outer membrane-enriched fractions (200 µg of protein) were solubilised in 20 µl of 125 mM ammonium bicarbonate solution containing 8 M urea before adding 4 µl of 12.5% (w/v) sodium deoxycholate (SDC). The samples were reduced with 10 µl of 60 mM DTT (final concentration of 10 mM) for 30 min and alkylated with 10 µl of 385 mM IAA (final concentration of 55 mM) for 30 min at room temperature in the dark. The concentration of urea in the samples was diluted from 8 M to 0.8 M by adding 700 µl of 27.5 mM

ammonium bicarbonate (final concentration of 25 mM) and further digested with trypsin (100 µl at 100 µg/ml in 25 mM ammonium bicarbonate), before incubating for 12-16 h at 37ºC. A volume of ethyl acetate was added equal to the volume of each sample to remove the SDC after trypsin digestion. The samples were acidified with 65 µl of 4% TFA (final concentration of 0.5% [v/v]) with shaking for 1 min and centrifuged at 15,700 x *g* for 2 min. The top layer of ethyl acetate and SDC was carefully removed before collecting digested peptides in the lower layer. The collected fractions were cleaned and desalted using a MEPS (Micro Extraction by Packed Sorbant, Presearch) syringe, and dried down to a 5 µl volume in a vacuum centrifuge. The samples were resolubilized in 20 µl of 2% acetronitrile and 0.1% formic acid prior to LC-ESI-Q/TOF MS analysis (**section 3.2.4.2.2**).

3.2.4.2 Proteomic identification by mass spectrometry

3.2.4.2.1 MALDI-TOF-TOF MS and data analysis

One microlitre of peptide solution (**section 3.2.4.1.1**) was mixed with an equivalent volume of saturated matrix solution (cyano-4-hydroxycinnamic acid, CHCA) on a MALDI-TOF target plate. Analysis was performed on an Applied Biosystem 4700 Proteomics Analyzer. The machine acquired MS/MS spectra from the eight most intense peak signals from the initial MS scan. Data generated from the MALDI-TOF-TOF mass spectrometer were used to perform searches of the eubacterial genome database using GPS Explorer software. Search parameters included peptide mass accuracy within 0.08 Da, one possible missed cleavage per peptide, variable methionine oxidation, and carbamidomethylation as a fixed modification. The significance of the identified proteins was judged based on protein scores greater than 79 ($p \leq 0.05$), observed pI and molecular mass, number of matched peptide masses, and percentage of sequence coverage. Unidentified samples were further analyzed by LC-ESI-Q/TOF MS.

3.2.4.2.2 LC-ESI-Q/TOF MS and data analysis

The peptide samples prepared by the gel-free methods (**section 3.2.4.1.2**) or not identified by MALDI-TOF-TOF MS (**section 3.2.4.1.1**) were analyzed by ESI-MS on a QSTAR XL Hybrid LC-MS/MS system as previously described by Bridges *et al.* (Bridges *et al.*, 2008). Data generated from the LC-ESI-Q/TOF mass spectrometer were analysed using Applied Biosystems Analyst QS version 1.1 and the automated Matrix Science MASCOT Daemon server version 2.1.06 (www.matrixscience.com). Proteins were identified using the MASCOT search engine against the eubacterial genome database. Variable methionine oxidation and carbamidomethylation as a fixed modification were used in the search options. The MS tolerance was set to 1.2 Da for MS and 0.6 Da for MS/MS analysis. The MASCOT program assigned a probability based MOWSE score to each protein. The identified proteins ($p \leq 0.05$) were significant if MOWSE scores were greater than 53.

3.2.5 Comparison of experimentally identified and confidently predicted OMPs

Proteins identified by both gel-based and gel-free proteomic techniques were integrated and compared to the list of confidently predicted putative OMPs described in **Chapter 2**. The OMPs identified from the eight representative isolates of *P. multocida* were compared in relation to the animal host of origin and disease syndrome. These analyses were performed using mySQL (Oracle), R package and Excel (Microsoft). Codon usage index analysis of the identified OMPs was computed by CodonW (http://codonW.sourceforge.net).

3.3 Results

3.3.1 OMPs prepared by different detergents

3.3.1.1 OMPs prepared by different concentrations of Sarkosyl

The use of different concentrations of Sarkosyl to prepare the outer membranes of the bovine isolate PM632 and the porcine isolate PM684 resulted in different OMP profiles (**Figure 3-2**). 0.1% Sarkosyl yielded a more complex OMP profile in both isolates. It was clearly seen from the polyacrylamide gel that the OmpA protein of isolate PM632 was reduced as the Sarkosyl concentration increases. Visible bands were cut from the gel and the proteins were identified by LC-MS/MS (**Figure 3-3 and 3-4**). Proteomic analyses of the OMP profiles prepared by using 0.1% Sarkosyl from both isolates showed that these profiles were contaminated with periplasmic proteins (such as SurA), inner membrane proteins (such as NAD(P) transhydrogenase, NqrA, NqrF, NqrC and YajC), cytoplasmic proteins (such as ribosomal proteins and elongation factors) (**Appendix Table 3-1**). Therefore, 0.1%, 1% and 2% Sarkosyl were not optimal for OMP preparation of *P. multocida* because of the high number of contaminated proteins (0.1% Sarkosyl) and the considerable loss of OmpA (1% and 2% Sarkosyl). This showed that 0.5% Sarkosyl was the optimum concentration.

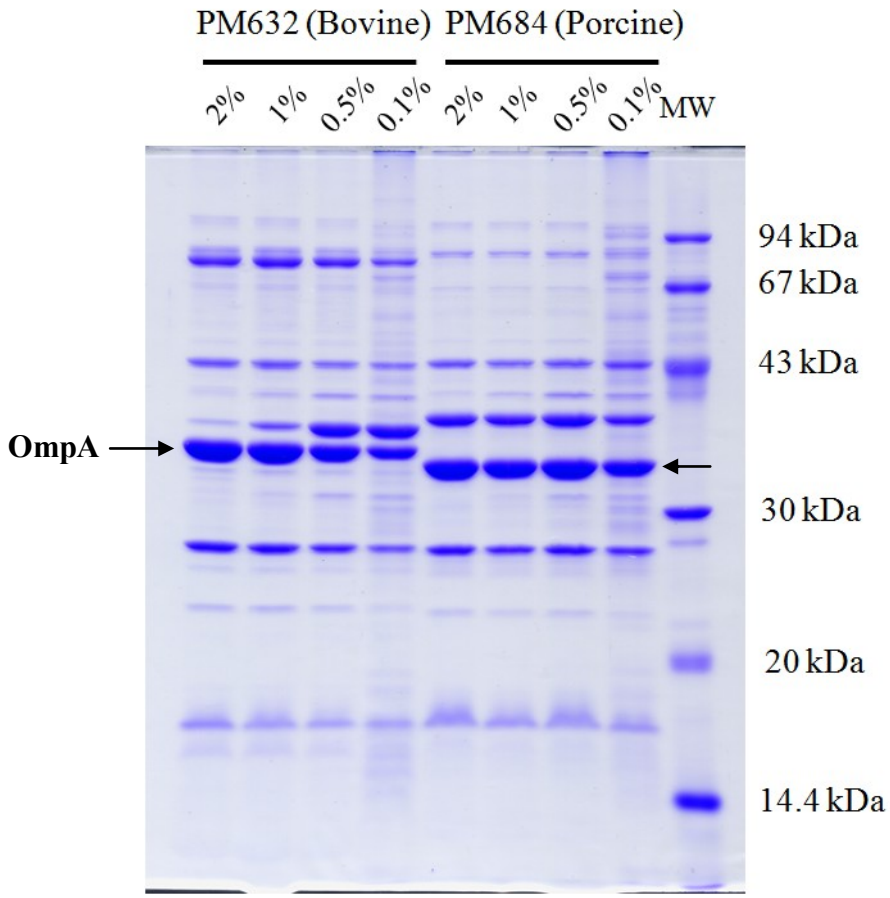


Figure 3-2. 1D SDS-polyacrylamide gel representing a comparison of the OMP profiles of isolates PM632 and PM684 prepared using different concentrations of Sarkosyl (0.1%, 0.5%, 1% and 2%). Arrows indicate the OmpA bands. The samples were prepared by in-gel trypsin digestion followed by LC-MS/MS.

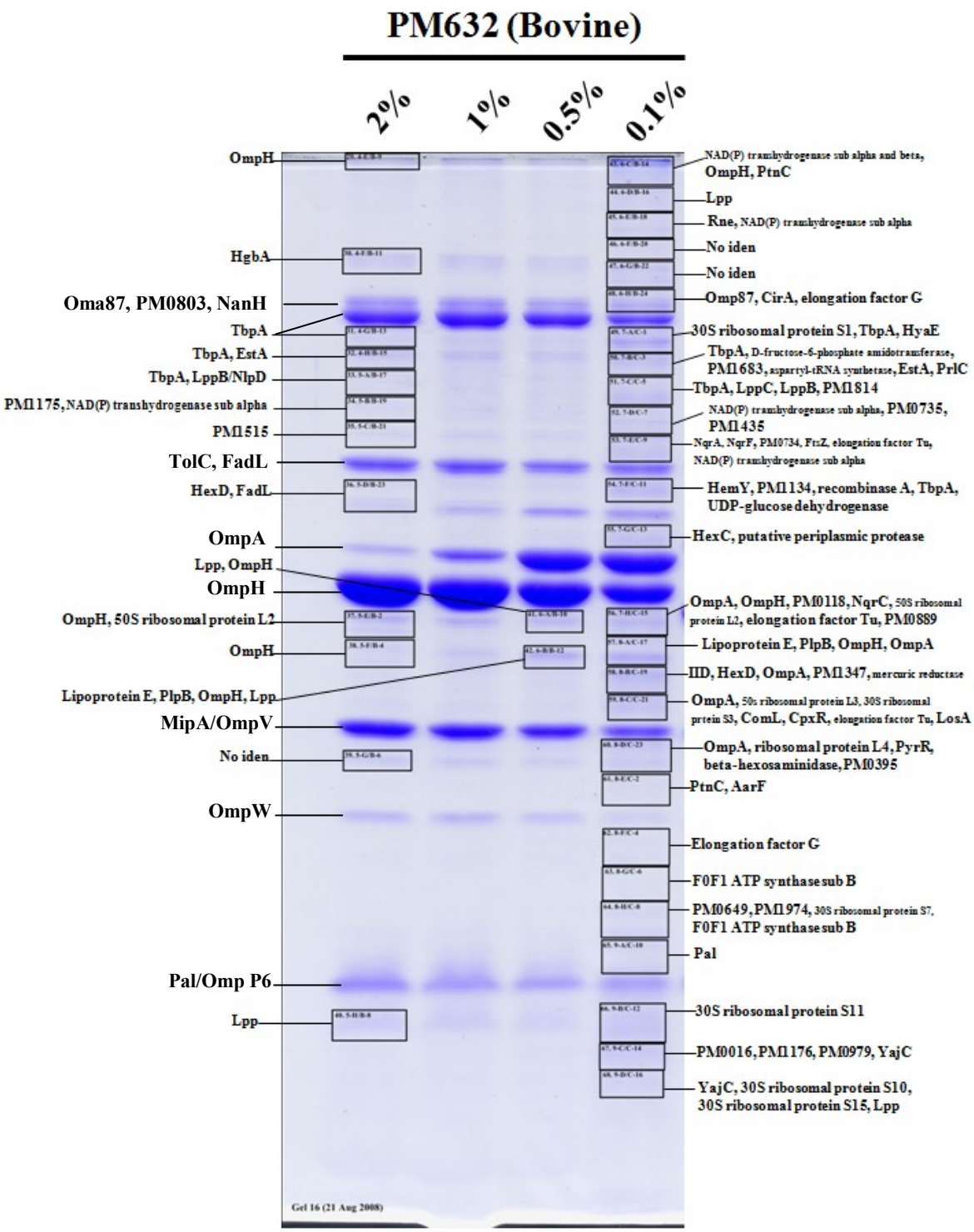


Figure 3-3. 1D-SDS-polyacrylamide gel showing excised bands from the OMP profiles of the bovine isolate PM632. The identified proteins corresponding to the cut bands are labelled.

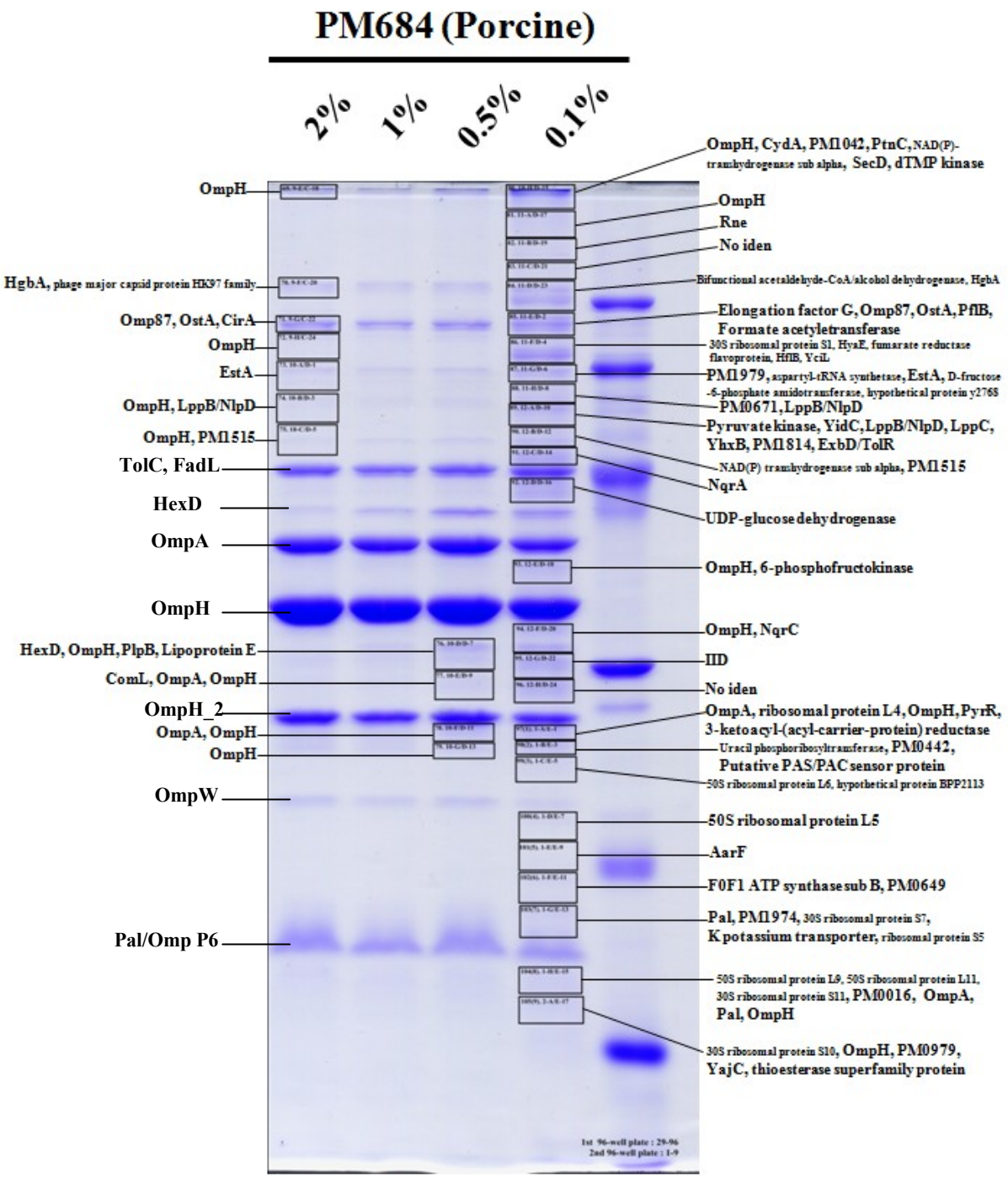


Figure 3-4. 1D-SDS-polyacrylamide gel showing excised bands from the OMP profiles of the porcine isolate PM684. The identified proteins corresponding to the cut bands are labelled.

3.3.1.2 OMPs prepared by different detergents

Seven different detergents were used to extract the outer membrane fractions of the isolate PM632 and compared with 0.5% Sarkosyl (**Figure 3-5**). Compared with 0.5% Sarkosyl, the use of 1% Sarkosyl with the addition of 7 mM EDTA caused considerable loss of the HexD (arrow number 1) and OmpA (arrow number 2) bands (**Figure 3-5**). The OMP profiles prepared with 2% Triton X-100, 1% Triton X-114, 0.5% CHAPS and 1% Octylglucoside were complex when compared to that of 0.5% Sarkosyl. The OMP profiles prepared by 1% Deoxycholate and carbonate wash yielded a moderate number of contaminant proteins. These bands were cut and analysed by LC-MS/MS. The results showed that the majority of the additional bands were contaminant proteins, especially ribosomal proteins (**Figures 3-6 and 3-7 and Appendix Table 3-2**). Therefore, in this study, the use of 0.5% Sarkosyl resulted in the cleanest and simplest OMP profile.

3.3.2 OMPs prepared by spheroplasting method

The spheroplasting method was examined in the bovine isolate PM632 and the porcine isolate PM684 (**Figure 3-8**). Four different fractions were obtained during the process: whole cells, spheroplast cells (containing cytoplasmic membrane), outer membrane fractions and precipitated supernatants after pelleting the outer membrane fractions. Clearly, the protein profiles obtained from the spheroplast cells, outer membrane fractions and precipitated supernatants were very similar to those of the whole cells in isolate PM684. These profiles were complex compared to the Sarkosyl-extracted OMP profiles. Therefore, this result showed that the spheroplasting method is less efficient in outer membrane extraction for *P. multocida* than in Sarkosyl extraction.

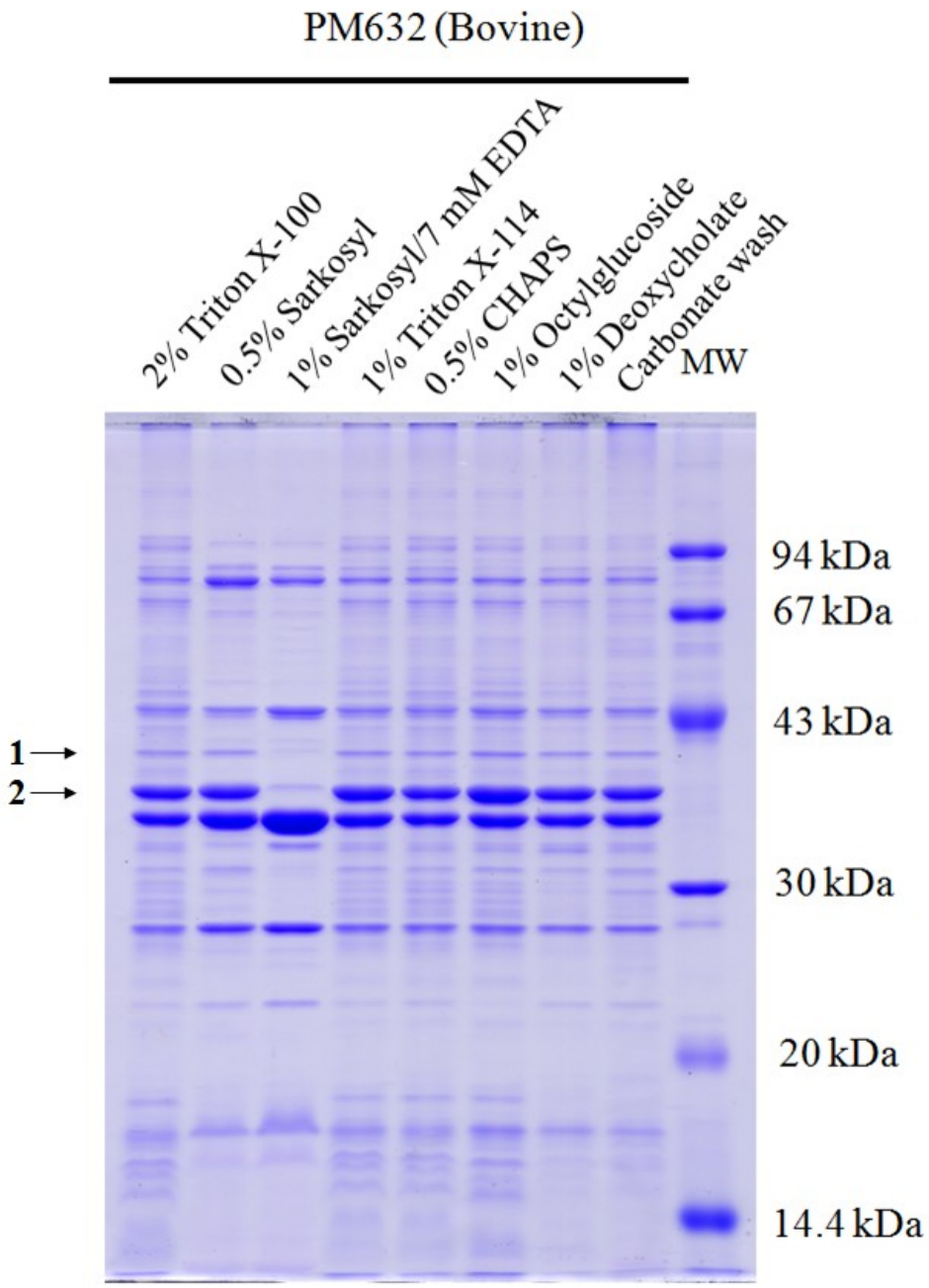


Figure 3-5. 1D SDS-polyacrylamide gel representing a comparison of the OMP profiles of bovine isolate PM632 prepared using different detergents. The samples were prepared by in-gel trypsin digestion followed by LC-MS/MS. Arrow number 1 represents HexD and arrow number 2 represents OmpA.

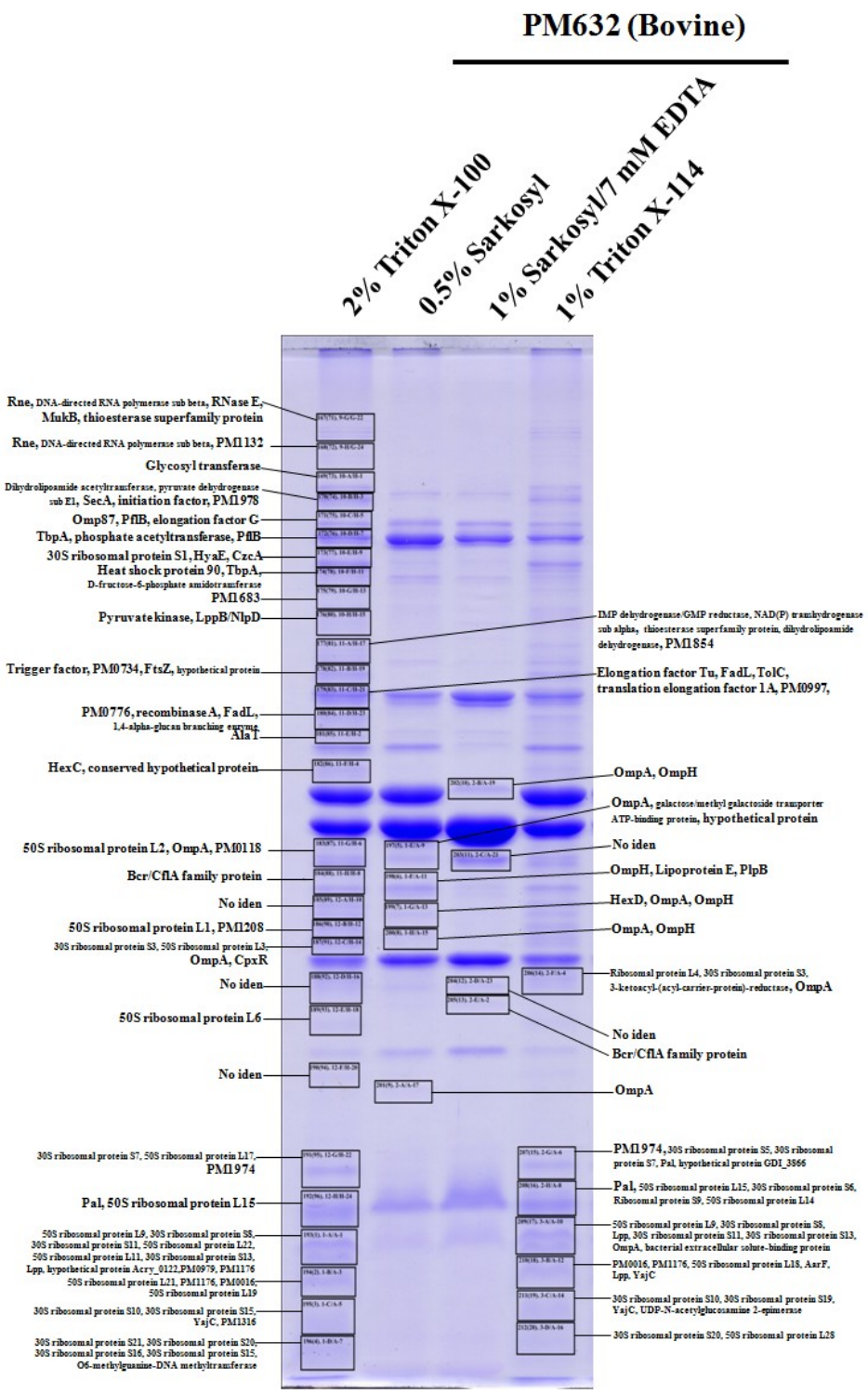


Figure 3-6. 1D-SDS-polyacrylamide gel showing excised bands from the OMP profiles of the bovine isolate PM632 prepared by 2% Triton X-100, 0.5% Sarkosyl, 1% Sarkosyl/7 mM EDTA and 1% Triton X-114. The identified proteins corresponding to the cut bands are labelled.

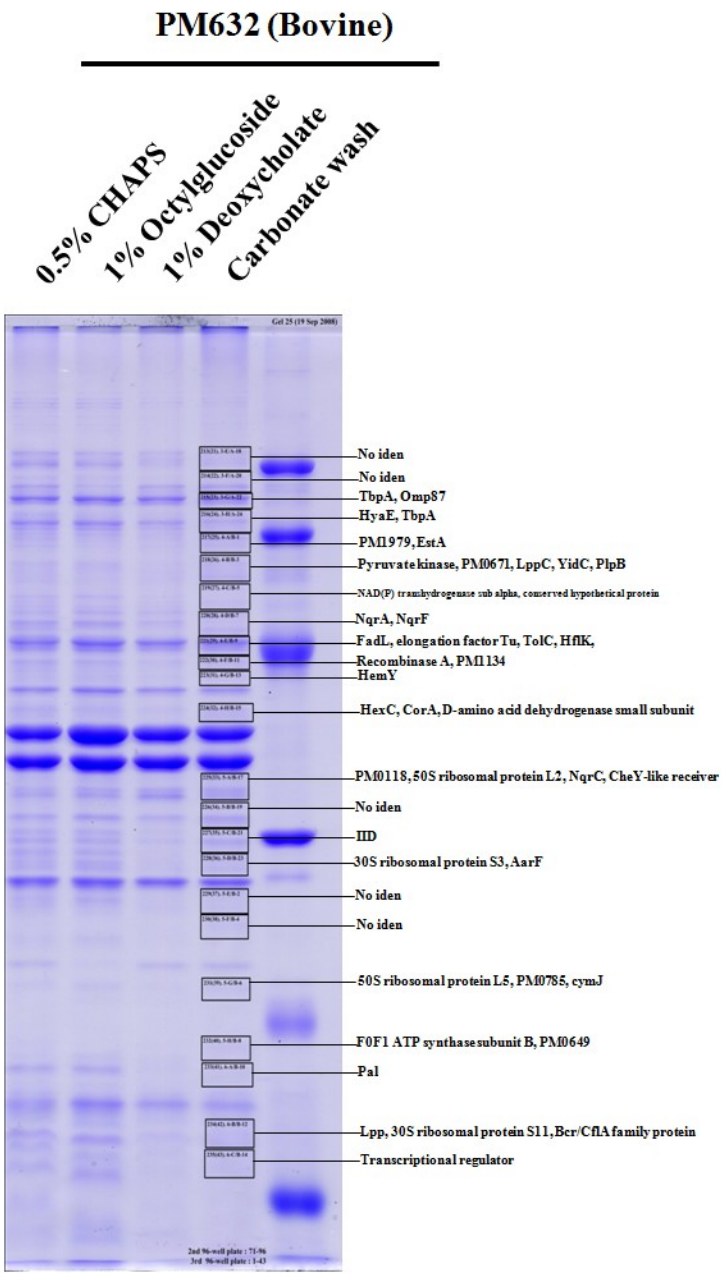


Figure 3-7. 1D-SDS-polyacrylamide gel showing excised bands from the OMP profiles of the bovine isolate PM632 prepared by 0.5% CHAPS, 1% Octylglucoside, 1% Deoxycholate and Carbonate wash. The identified proteins corresponding to the cut bands are labelled.

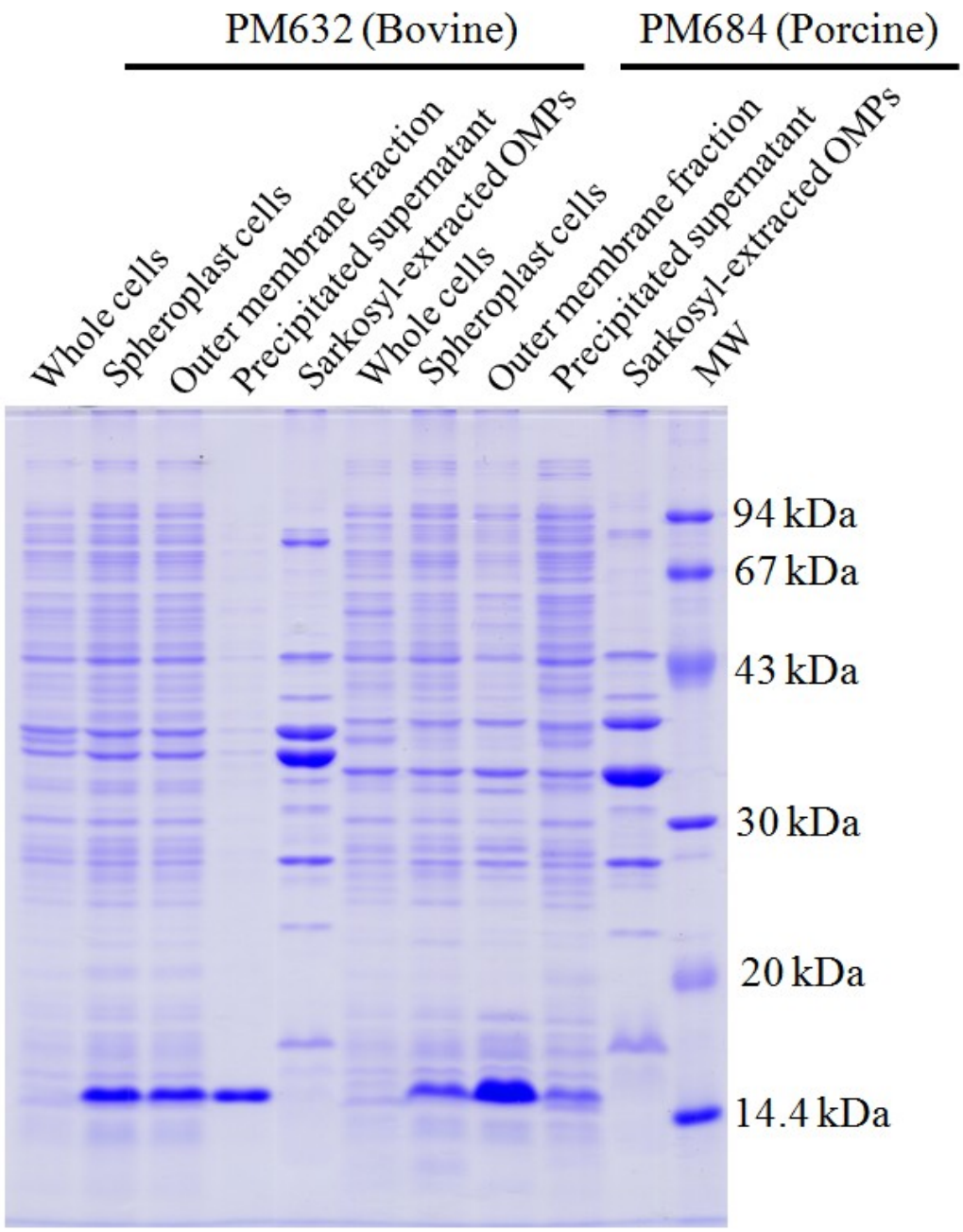


Figure 3-8. 1D SDS-polyacrylamide gel representing different fractions (whole cells, spheroplast cells, outer membrane fractions and precipitated supernatants) obtained during the spheroplsting method of two isolates of *P. multocida* compared with those prepared by Sarkosyl extraction.

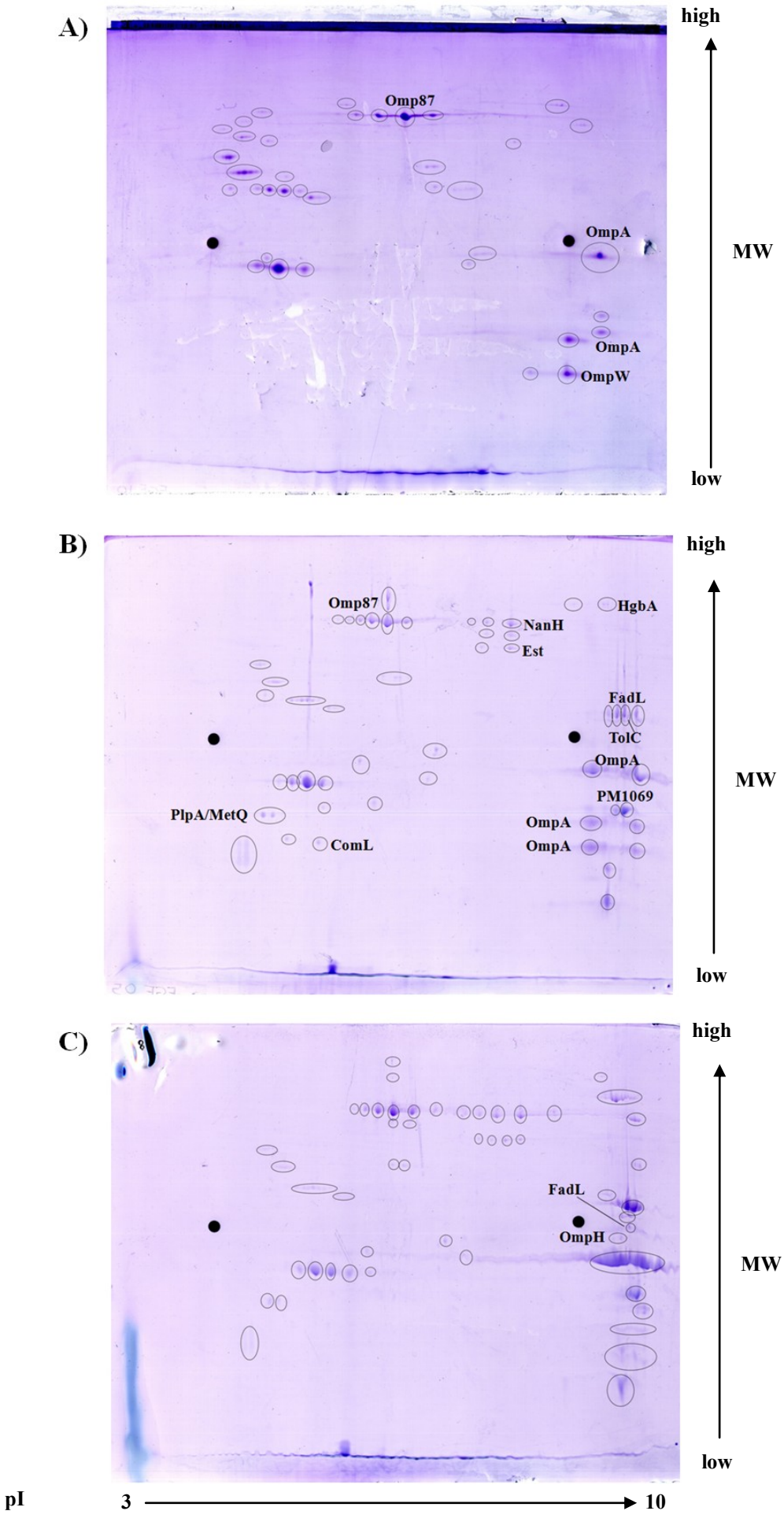
3.3.3 2D SDS-PAGE of the outer membrane proteome of P. multocida

Three different 2D SDS-PAGE protocols were used to separate the outer membrane proteome of the bovine isolate PM632 of *P. multocida*. The major differences of these three methods were in the composition of the rehydration buffer: the first method used CHAPS in the buffer; the second method boiled the sample before resuspending in buffer containing ASB-14; and the third method used a combination of CHAPS and ASB-14 in the buffer. The 2D gels prepared by these three protocols were very similar (**Figure 3-9**). The presence of ASB-14 in the rehydration buffer of the second and third methods (**Figures 9B and 9C**) resulted in the presence of more protein spots on the gels. All visible protein spots from the gels of these three methods were cut out and analyzed by gel-based method followed by LC-MS/MS and 11 OMPs were identified. These included Oma87, OmpA, OmpW, NanH, HgbA, FadL, EstA, PlpA/MetQ, ComL, TolC and OmpH. All of the identified proteins were transmembrane β-barrel proteins; no outer membrane lipoproteins were detected. Certain OMPs such as Oma87 and OmpA showed multiple spots on the gels, indicating possible multiple isoforms of these proteins or their modification. These results showed that fewer proteins were identified in 2D SDS-PAGE than in 1D SDS-PAGE (**section 3.3.4.1**).

3.3.4 Comparative proteomic analyses of the OMPs prepared by Sarkosyl extraction

3.3.4.1 Identification of OMPs by gel-based and gel-free proteomic approaches

Extraction of the outer membrane fractions by different concentrations (0.1%, 0.5%, 1% and 2%) of Sarkosyl revealed that 0.5% Sarkosyl was the optimal concentration for this


A)
high
Omp87
OmpA
MW
OmpA
OmpW
low
B)
high
Omp87
HgbA
NanH
Est
FadL
TolC
OmpA
MW
PM1069
PlpA/MetQ
OmpA
ComL
OmpA
low
C)
high
FadL
OmpH
MW
low
pI
3
10

bacterium (**Figure 3-2**). A "standard" 1-D SDS-polyacrylamide gel of the OMP profiles (0.5% Sarkosyl extraction) of the eight representative isolates is presented in **Figure 3-10** and shows the extensive variation between isolates in the OMP profiles in terms of molecular mass and expression. The abundantly-expressed bands, 1 and 3, represent the major OMPs, OmpH and OmpA, respectively, and these varied in molecular mass across all isolates. The high level of expression of the major OMPs was consistent with their predicted high codon adaptation index (CAI) values, 0.383 for OmpA and 0.372 for OmpH. High and low molecular mass forms of OmpA were also identified and this is consistent with the heat-modifiable properties of this protein. Certain minor proteins were present in all isolates and exhibited varying degrees of molecular mass variation (e.g., proteins 5 and 17). In other cases, proteins appeared to be present in some but not all isolates (e.g., protein 26 was present only in the two bovine and a single ovine isolate). Positive identifications were obtained for 44 OMPs by the gel-based approach **(Figure 3-10, Table 3-3, Appendix Tables 3-3 and 3-4)**. Of these, twenty-seven had transmembrane β-barrel structures and 17 were lipoproteins. Faintly staining bands in the area between 43 kDa and 67 kDa (proteins 9, 13, 14, 26, 27, 29, 37, 52) and in the 14 kDa (proteins 18, 20, 31) region **(Figure 3-10)** represented multiple proteins that were expressed at low level. Some apparently single bands (e.g., proteins 6 and 11, and 15 and 19) comprised two or more proteins.

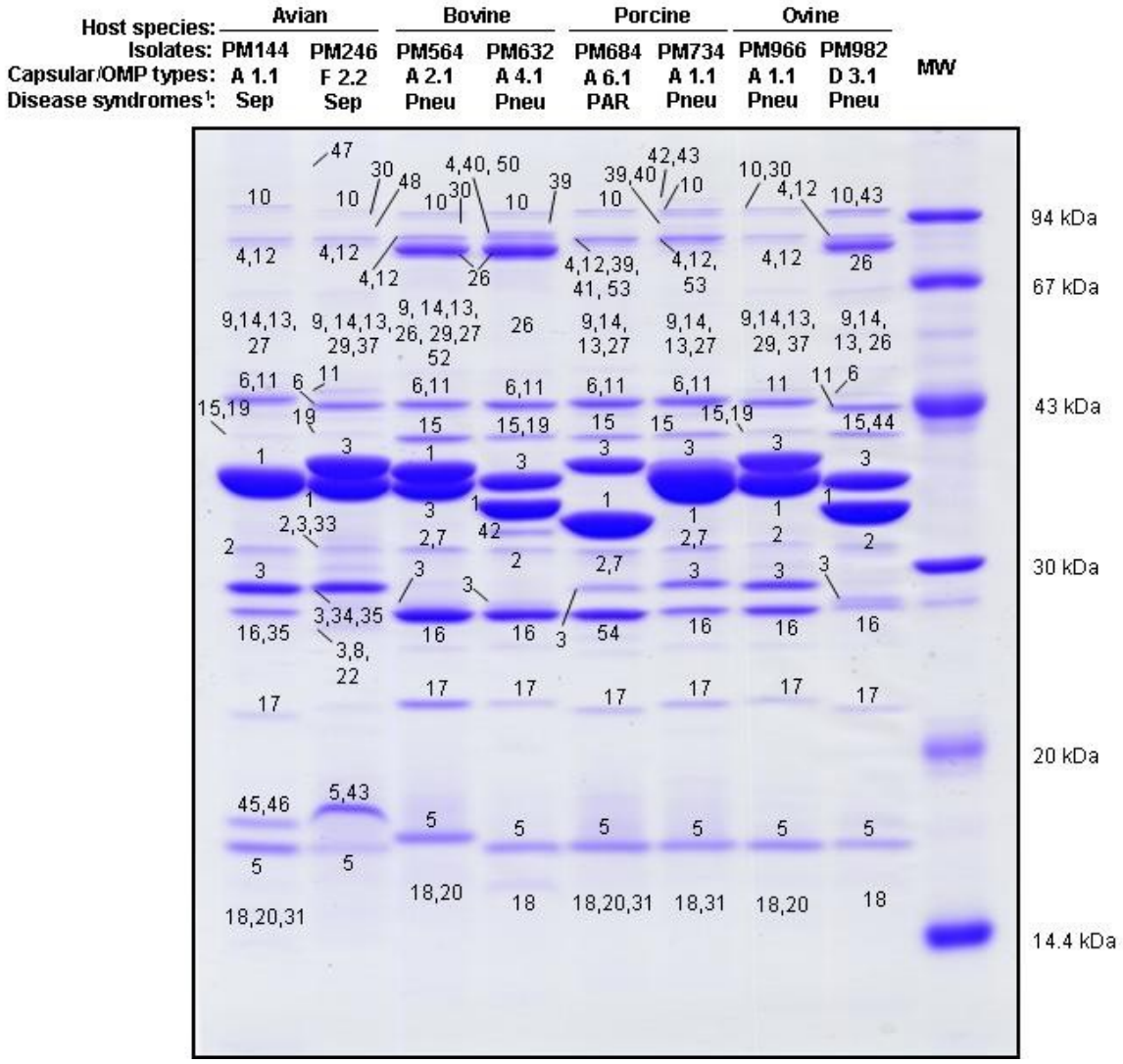


Figure 3-10. 1-D 12% SDS-polyacrylamine gel representing the gel-based proteomic identification of the OMPs from eight representative isolates of *P. multocida* recovered from different animal hosts and disease syndromes. The outer membrane-enriched fractions were prepared by 0.5% Sarkosyl extraction, separated by 1-D SDS-PAGE and stained with Coomassie brilliant blue. Twenty micrograms of protein were loaded per lane and the gel was run until the dye-front reached the end of the gel. Labelled numbers on the gel correspond to numbers of proteins in Table 2. Molecular weight markers (MW) are shown on the right.

[1] Disease symdromes: Sep, septicaemia; Pneu, pneumonia; PAR, progressive atrophic rhinitis

Table 3-3. Proteins identified from the outer membrane-enriched fractions of the eight representative isolates of *P. multocida* obtained from four different animal hosts using a combination of gel-based and gel-free proteomic techniques.

No.	Protein ID[a]	Protein name	Function	Isolates of different animal hosts[b]								CAI[C]
				Avian		Bovine		Porcine		Ovine		
				PM144	PM246	PM564	PM632	PM684	PM734	PM966	PM982	
1	PM0388	OmpH_1	Porin, ion transport activity	+	+	+	+	+	+	+	+	0.372
2	PM1730	PlpA/MetQ	Amino acid transport	+	+	+	+	+	+	+	+	0.308
3	PM0786	OmpA	Outer membrane integrity	+	+	+	+	+	+	+	+	0.383
4	PM1992	Oma87	OM biogenesis and surface antigen	+	+	+	+	+	+	+	+	0.256
5	PM0966	Pal/Omp P6	Envelope integrity/link outer membrane to peptidoglycan	+	+	+	+	+	+	+	+	0.388
6	PM0527	OM efflux protein TolC	Protein secretion and transport activity	+	+	+	+	+[1]	+	+[2]	+	0.204
7	PM1064	Lipoprotein E/OmpP4	Acid phosphatase activity and utilization of NAD, NADP	+[2]	+[2]	+	+[2]	+	+	+[2]	+[2]	0.319
8	PM1501	VacJ	Promoting spread of bacteria through tissues	+[2]	+	+[2]	+[2]	+[2]	+[2]	+[2]	+[2]	0.249
9	PM0076	EstA	Lipid metabolism/hydrolase activity, acts on ester bonds	+[1]	+	+[1]	+[2]	+	+[1]	+	+	0.232
10	PM0300	HgbA	Transport, haemoglobin and haemoglobin-haptoglobin uptake	+	+	+[1]	+	+	+	+	+	0.228
11	PM1069	FadL	Transport of hydrophobic compounds	+[1]	+[1]	+	+[1]	+[1]	+	+	+	0.286
12	PM1600	LptD/Imp/OstA	LPS assembly/response to organic substance	+[1]	+	+[1]	+[2]	+[1]	+	+	+[1]	0.243

[a]Uniprot protein ids

[b]Two proteomic methods were compared; '+[1]' = proteins identified by gel-based method; '+[2]' = proteins identified by gel-free method; '+' = proteins identified by both methods; '-' = no identification.

[c]Codon adaptation index

Table 3-3. (continued)

No.	Protein ID	Protein name	Function	Isolates of different animal hosts								CAI
				Avian		Bovine		Porcine		Ovine		
				PM144	PM246	PM564	PM632	PM684	PM734	PM1966	PM982	
13	PM1809	YtfM/Omp85 family protein	Unknown	+	+	+	-	+	+	+	+	0.234
14	PM1614	LppB/NlpD	Cell wall catabolic process and proteolysis	+	+	+[1]	-	+	+	+[1]	+[1]	0.244
15	PM0778	HexD	Capsular polysaccharide transport	+	-	+	+	+	+	+	+	0.208
16	PM0998	MipA/OmpV family protein	Outer membrane biogenesis/MltA-interacting protein	+	-	+	+	+[2]	+	+	+[1]	0.215
17	PM0331	OmpW	Transport small hydrophobic molecules	+[1]	-	+[1]	+[1]	+[1]	+[1]	+	+[1]	0.24
18	PM0554	Lpp/Pcp	Unknown, surface-exposed	+[1]	-	+[1]	+[1]	+[1]	+[1]	+[1]	+[1]	0.251
19	PM1444	GlpQ	Glycerol metabolic process, lipid metabolic process	+	+	+[2]	+	+[2]	+[2]	+	-	0.29
20	PM0016	Lipoprotein	Unknown	+	+[2]	+	+[2]	+	+[2]	+	-	0.223
21	PM0442	Lipoprotein	Unknown	+[2]	+[2]	+[2]	+[2]	-	+[2]	+[2]	-	0.49
22	PM1720	ComL	DNA uptake/outer membrane biogenesis	+[2]	+[1]	+[2]	-	-	+[2]	+[2]	+[2]	0.265
23	PM1050	NlpB	Outer membrane biogenesis/insertion of OMPs	+[2]	-	-	+[2]	+[2]	+[2]	+[2]	+[2]	0.23
24	PM1826	Hypothetical protein	Unknown	+[2]	+[2]	+[2]	-	-	+[2]	+[2]	-	0.257
25	PM1827	Hypothetical protein	Unknown	+[2]	+[2]	-	-	+[2]	+[2]	+[2]	-	0.239
26	-	TbpA	Transferrin receptor and transport activities	-	-	+	+	+[2]	-	-	+	0.211
27	PM1515	Conserved hypothetical protein	Unknown	+[1]	-	+[1]	-	+[1]	+	-	-	0.213
28	PM1426	Phospholipase A/OmpLA	Lipid metabolic process/maintain asymmetry of the OM	-	+[2]	-	+[2]	-	+[2]	-	-	0.266

Table 3-3. (continued)

No.	Protein ID	Protein name	Function	Isolates of different animal hosts								CAI
				Avian		Bovine		Porcine		Ovine		
				PM144	PM246	PM564	PM632	PM684	PM734	PM966	PM982	
29	PM0056	LspB_1	Two-partner secretion/secretion of filamentous haemagglutinin	-	+[1]	+	-	-	-	+[1]	-	0.214
30	PM0266	Mce/PqiB	Unknown	-	+[1]	+[1]	-	-	-	+[1]	-	0.22
31	PM1886	SmpA	Maintaining envelope integrity and beta barrel OMP assembly	+[1]	-	-	-	+[1]	+[1]	-	-	0.234
32	PM1215	RlpB	LPS assembly/outer membrane biogenesis	-	+[2]	-	-	+[2]	-	-	-	0.225
33	PM0586	Plp4	Unknown	+[2]	+	-	-	-	-	+[2]	-	0.308
34	PM0852	RcpA	Protein secretion/Flp pilus biogenesis	+[2]	+	-	-	-	-	+[2]	-	0.199
35	PM0846	TadD	Protein secretion/binding/assembly and transport of Flp pili	+	+	-	-	-	-	+[2]	-	0.19
36	PM0853	RcpC	Tight adherence and fibril production	-	+[2]	-	-	-	-	+[2]	-	0.207
37	PM0058	LspB_2	Two-partner secretion/secretion of filamentous haemagglutinin	-	+[1]	-	-	-	-	+[1]	-	0.214
38	PM0708	Slp	Starvation-inducible lipoprotein	-	-	-	-	-	-	-	+[2]	0.202
39	PM1622	HasR	Haem receptor and transport activities	-	-	-	+[1]	+[1]	+	-	-	0.193
40	PM0803	TonB-dependent receptor	Receptor and transport activities	-	-	-	+	-	+	-	-	0.225
41	PM1428	TonB-dependent receptor	Receptor and transport activities	-	-	-	-	-	+	+[2]	-	0.233
42	PM0336	HgbB	Transport, haemoglobin and hemoglobin-haptoglobin uptake	-	-	-	-	-	+[1]	+$_2$	-	0.265

Table 3-3. (continued)

No.	Protein ID	Protein name	Function	Isolates of different animal hosts								CAI
				Avian		Bovine		Porcine		Ovine		
				PM144	PM246	PM564	PM632	PM684	PM734	PM966	PM982	
43	PM0337	HgbB	Transport, haemoglobin and hemoglobin-haptoglobin uptake	-	-	-	-	-	+[1]	-	+[1]	0.266
44	PM1134	Hypothetical protein	Protein secretion/protein transport activity	-	-	-	-	-	+[1]	-	+[1]	0.232
45	PM1025	Opa	Porin activity and adherence	+[1]	+[1]	-	-	-	-	-	-	0.259
46	PM1077	HlpB	Unknown	+	-	-	-	-	-	-	-	0.229
47	PM1808	OmpL41/YtfN-like protein	Bacterial morphogenesis	-	+[1]	-	-	-	-	-	-	0.206
48	PM1717	OM autotransporter	Unknown	-	+[1]	-	-	-	-	-	-	0.226
49	PM0851	RcpB	Unknown	-	+[2]	-	-	-	-	-	-	0.233
50	PM0663	NanH	Exo-alpha-sialidase/produce free sialic acid as energy and carbon sources	-	-	-	+	-	-	-	-	0.197
51	PM0395	YccT	Unknown	-	-	-	+[2]	-	-	-	-	0.271
52	PM0646	LppC	Unknown	-	-	+[1]	-	-	-	-	-	0.235
53	PM0576	HemR	Haem receptor and transport activities	-	-	-	-	+[1]	+[1]	-	-	0.179
54	PM0389	OmpH_2	Porin, ion transport activity	-	-	-	-	+[1]	-	-	-	0.305
			Total	**32**	**33**	**28**	**26**	**29**	**35**	**34**	**24**	

A combination of the methanol-aided and phase transfer surfactant-aided trypsin digestion gel-free approaches identified 43 OMPs in outer membrane preparations **(Table 3-3, Appendix table 3-5)**. These significantly-detected OMPs were those that had MOWSE scores greater than 53 and were identified by at least one peptide. These included 24 transmembrane β-barrel proteins and 19 lipoproteins**.** Integration of the results from the gel-based and gel-free proteomic approaches resulted in the identification of 54 OMPs **(Table 3-3).** Fifty-one of these proteins were predicted to be representative of the OM proteomes of avian strain Pm70 (52%) or porcine strain 3480 (48%). Three proteins (TbpA, RcpB, and hypothetical protein PM1134) were absent from our list of confidently predicted OMPs. Transferrin binding protein A, TbpA, is not present in the genomes of avian strain Pm70 and porcine strain 3480, while RcpB and PM1134 did not pass the filtering criteria and were screened out from both genomes. However, information from the literature and protein database searches indicated that these two proteins are indeed potential OMPs. RcpB is part of the *tad* locus (which is involved in biofilm formation) although the precise function of the protein is not known; PM1134 is a hypothetical protein that has protein secretion and transport activities (**Table 3-3**).

Thirty-two OMPs were identified by both gel-based and gel-free methods; 19 of these had β-barrel structures and 13 were lipoproteins (**Figure 3-11A).** The gel-based approach identified 12 unique OMPs (SmpA, LspB_2, HgbB, OmpH_2, Opa, uncharacterized autotransporter PM1717, PqiB, OmpL41-like protein, HemR, LppC, Lpp/Pcp, and hypothetical protein PM1134); the majority (9/12) of these were β-barrel proteins (**Figure 3-11A**). Similarly, 10 OMPs (Slp, NlpB, RlpB, RcpC, RcpB, OmpLA, YccT, uncharacterized lipoprotein PM0442, PM1826, and PM1827) were uniquely identified by the gel-free approach; in contrast, the majority (7/10) of these were lipoproteins (**Figure 3-11A**).

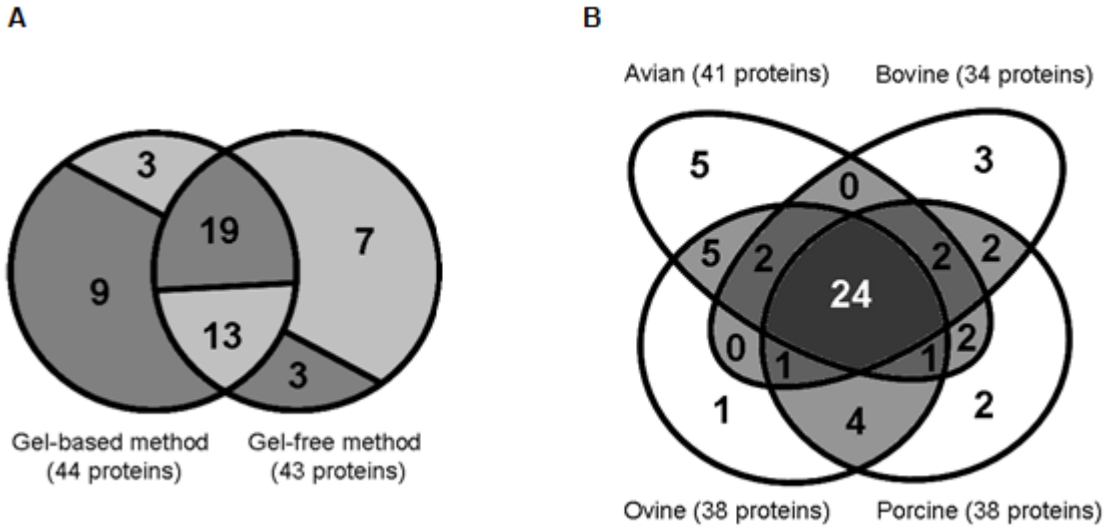


Figure 3-11. A) A comparison of the number of OMPs identified using the gel-based and gel-free methods. The areas shaded in dark gray represent transmembrane β-barrel proteins and the light gray areas represent lipoproteins and proteins of unknown structure. Numbers indicate the number of proteins. The total number of OMPs identified by each method is shown in parentheses. B) A comparison of the OMPs identified in isolates from four different animal hosts (avian, bovine, porcine and ovine). The OMPs identified in four, three, two and one animal hosts are shown in dark gray-, medium dark gray-, light gray-shaded and non-shaded areas, respectively. Numbers indicate the number of proteins. The total number of OMPs associated with each host is shown in parentheses.

3.3.4.2 Comparison of OMPs associated with different host species and disease syndromes

The number of proteins identified in each of the eight isolates varied from 24 (isolate PM982) to 35 (isolate PM734) (average of 30 $\pm$ 4 proteins) **(Table 3-3)**. The greatest variation in the number of identified proteins between pairs of isolates was observed in the porcine (29 and 35) and ovine (24 and 34) isolates. Comparison of the OMPs identified in isolates from each of the four animal hosts **(Figure 3-11B, Table 3-3)** revealed two major groups of proteins which we defined as core and host-associated proteins. Core proteins are those associated with either one or two isolates from all four animal hosts; host-associated proteins are those associated with either one or two isolates from one to three animal hosts. Twenty-four core proteins were identified that were associated with isolates from all four animal hosts (**Figures 3-11B and 3-12**). Proteins OmpH_1, PlpA/MetQ, OmpA, Oma87, Pal, TolC, Lipoprotein E/OmpP4, VacJ, EstA, HgbA, FadL, and LptD/Imp were identified in both isolates from each animal host, whereas YtfM/Omp85 family protein, LppB/NlpD, HexD, MipA, OmpW, Lpp/Pcp, GlpQ, lipoproteins PM0016 and PM0442, ComL, NlpB, and hypothetical protein PM1826 were identified in at least one isolate from each animal host **(Figure 3-11B, Table 3-3)**. The functions of these proteins include outer membrane biogenesis and integrity (OmpA, Oma87, Pal, LptD/Imp, LppB/NlpD, MipA, NlpB), transport and receptor (OmpH_1, PlpA/MetQ, TolC, HgbA, FadL, HexD, OmpW, ComL), enzymatic activity (Lipoprotein E/OmpP4, EstA, GlpQ) and unknown (VacJ, YtfM/Omp85 family protein, Lpp/Pcp, PM0016, PM0442, PM1826).

Six core proteins (lipoprotein PM0442, Pal, OmpA, OmpH_1, Lipoprotein E/OmpP4 and PlpA/MetQ) and two host-associated proteins (Plp4 and OmpH_2) had the highest predicted CAI scores (> 0.30) indicating high expression levels (**Figure 3-13**). Forty-one

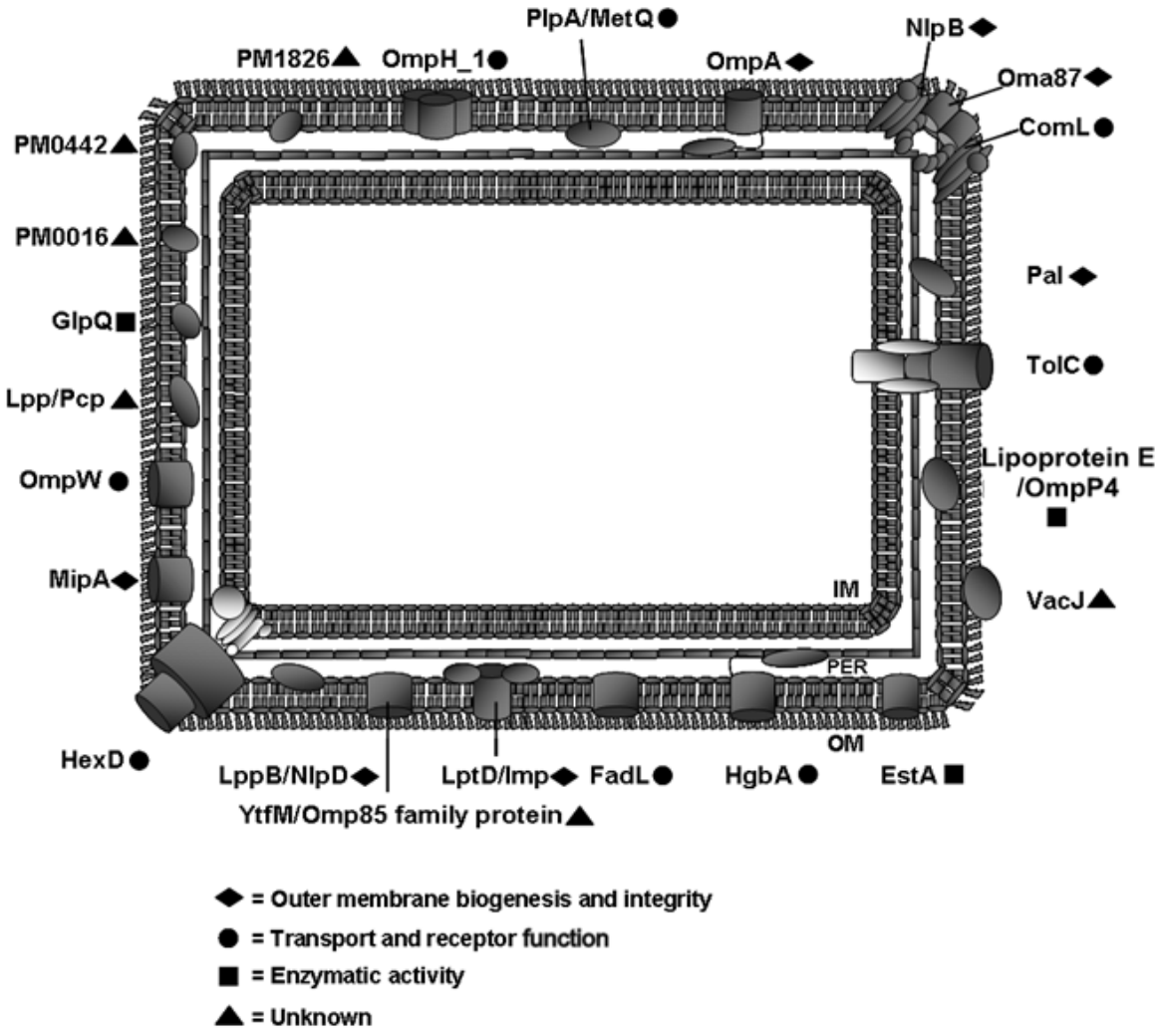


Figure 3-12. Core proteins associated with all four animal hosts. These proteins have roles in outer membrane biogenesis and integrity, transport and receptor function, and enzymatic activity. Cylindrical shapes represent topology of transmembrane proteins while oval shapes represent lipoproteins or periplasmic/plug domains of OmpA and HgbA. The order of proteins (starting with OmpH_1) corresponds to that of **Table 3-3**, except NlpB, Oma87 and ComL which form an outer membrane protein assembly complex.

proteins had CAI values between 0.2 and 0.3 (**Figure 3-13**) and five proteins (RcpA, NanH, HasR, TadD and HemR) had CAI values below 0.2.

Host-associated proteins **(Figure 3-11B, Table 3-4)** included 11 proteins that were restricted to isolates from only one host (Opa, RcpB, HlpB, OmpL41-like protein, autotransporter PM1717, NanH, YccT, LppC, HemR, OmpH_2 and Slp), 13 proteins that were identified in isolates from two animal hosts (SmpA, RlpB, LspB_2, RcpA, RcpC, TadD, Plp4, HasR, TonB-dependent receptor proteins PM0803 and PM1428, two HgbB and hypothetical protein PM1134) and six proteins that were identified in isolates from three hosts (phospholipase A/OmpLA, hypothetical protein PM1515, LspB_1, Mce/PqiB, hypothetical protein PM1827 and TbpA). Certain of these proteins are well characterized whereas the functions of others have been derived from their homologues in other bacteria. RcpA, RcpC and TadD have a role in the formation of the Flp pilus and were expressed only in the avian and ovine isolates. SmpA and RlpB are involved in outer membrane biogenesis and integrity and were identified in avian and porcine isolates. OmpLA is an outer membrane phospholipase and was identified in avian, bovine and porcine isolates. LspB_1 and LspB_2 function in the secretion of filamentous hemagglutinin and were found in avian, bovine and ovine and in avian and ovine isolates, respectively; neither occurred in porcine isolates. The transferrin, heme and hemoglobin receptors (TbpA, HasR, HgbB) and uncharacterized TonB-dependent proteins (PM0803 and PM1428) were absent from the avian isolates. TbpA was identified in bovine, porcine and ovine isolates; HasR and PM0803 were found in bovine and porcine isolates; and HgbB and PM1428 occurred in porcine and ovine isolates. Opa, RcpB, HlpB, OmpL41-like protein and autotransporter PM1717 were identified only in the avian isolates, whereas NanH, YccT and LppC were associated only with bovine isolates, HemR and OmpH_2 only with porcine isolates and Slp only with ovine isolates **(Table 3-4)**.

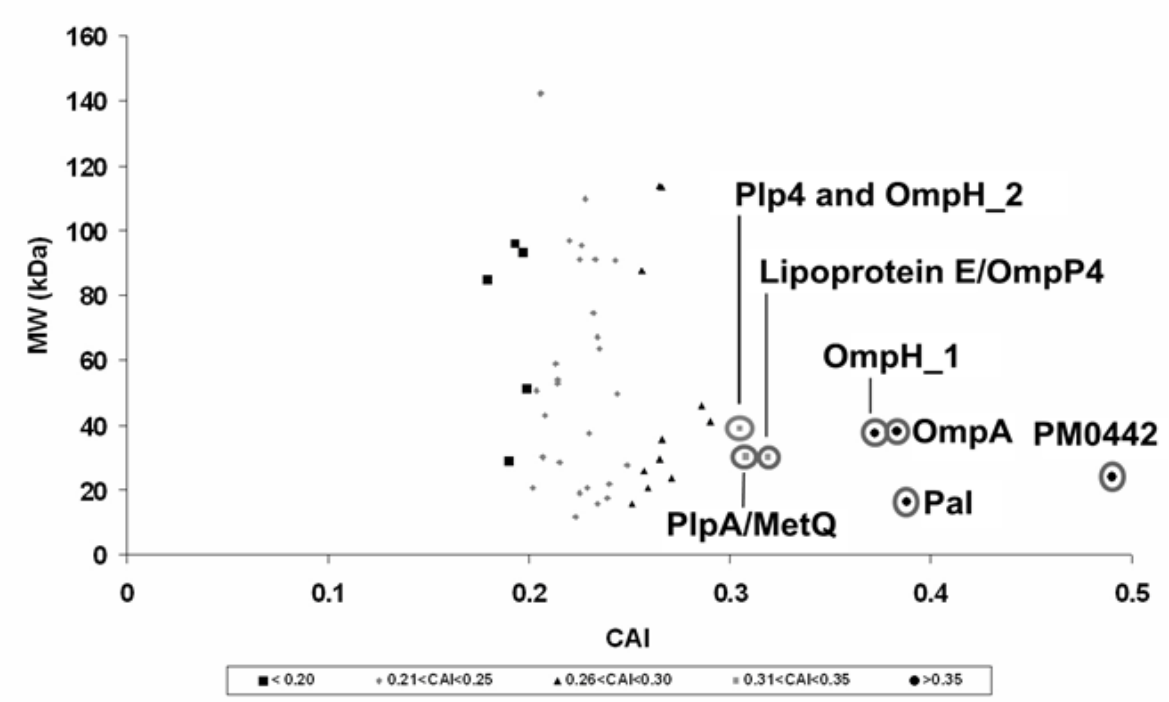


Figure 3-13. Graph showing a comparison between molecular mass (MW, y axis) and predicted CAI score (x axis) of the identified OMPs. The CAI scores were grouped into five categories: <0.20, 0.21 – 0.25, 0.26 – 0.30, 0.31 – 0.35, and >0.35. The higher CAI score means the probability of higher level of expression of that particular protein. The OMPs which have high CAI scores are labelled.

Table 3-4. Association of host-associated OMPs with their avian, bovine, porcine and ovine hosts. These proteins have roles in outer membrane biogenesis and integrity, receptor-mediated transport, transport, adherence, and enzymatic activity or have unknown functions. + indicates that the protein was identified in one or more isolates from that particular host; - indicates that the protein was not identified.

No	Protein ID	Protein name	Function	Animal host of origin			
				Avian	Bovine	Porcine	Ovine
1. OMPs identified in one host							
1	PM1025	Opa	Adherence	+	-	-	-
2	PM0851	RcpB	Adherence	+	-	-	-
3	PM1077	HlpB	Unknown	+	-	-	-
4	PM1808	OmpL41/YtfN-like protein	Unknown	+	-	-	-
5	PM1717	OM autotransporter	Unknown	+	-	-	-
6	PM0663	NanH	Enzymatic activity	-	+	-	-
7	PM0395	YccT	Unknown	-	+	-	-
8	PM0646	LppC	Unknown	-	+	-	-
9	PM0576	HemR	Receptor-mediated transport	-	-	+	-
10	PM0389	OmpH_2	Transport	-	-	+	-
11	PM0708	Slp	Unknown	-	-	-	+
2. OMPs identified in two hosts							
1	PM1886	SmpA	OM biogenesis and integrity	+	-	+	-
2	PM1215	RlpB	OM biogenesis and integrity	+	-	+	-
3	PM0058	LspB_2	Receptor-mediated transport	+	-	-	+
4	PM0852	RcpA	Adherence	+	-	-	+

Table 3-4. (continued)

No	Protein ID	Protein name	Function	Animal host of origin			
				Avian	**Bovine**	**Porcine**	**Ovine**
5	PM0853	RcpC	Adherence	+	-	-	+
6	PM0846	TadD	Adherence	+	-	-	+
7	PM0586	Plp4	Unknown	+	-	-	+
8	PM1622	HasR	Receptor-mediated transport	-	+	+	-
9	PM0803	TonB-dependent receptor	Receptor-mediated transport	-	+	+	-
10	PM1428	TonB-dependent receptor	Receptor-mediated transport	-	-	+	+
11	PM0336	HgbB	Receptor-mediated transport	-	-	+	+
12	PM0337	HgbB	Receptor-mediated transport	-	-	+	+
13	PM1134	Hypothetical protein	Receptor-mediated transport	-	-	+	+
3. OMPs identified in three hosts							
1	PM1426	Phospholipase A/OmpLA	Enzymatic activity	+	+	+	-
2	PM1515	Conserved hypothetical protein	Unknown	+	+	+	-
3	PM0056	LspB_1	Receptor-mediated transport	+	+	-	+
4	PM0266	Mce/PqiB	Unknown	+	+	-	+
5	PM1827	Hypothetical protein	Unknown	+	-	+	+
6	-	TbpA	Receptor-mediated transport	-	+	+	+

The OMPs identified in isolates associated with different disease status were also compared (**Figure 3-14**). As expected, the 24 core proteins were associated with isolates recovered from cases of septicaemia, pneumonia, and PAR. Eleven OMPs were shared between isolates causing septicaemia and pneumonia. In contrast, the single isolate associated with PAR shared only one and three OMPs with isolates responsible for septicaemia and pneumonia, respectively. Nine proteins were identified only in isolates associated with pneumonia, five proteins were identified only in isolates recovered from cases of septicaemia and a single protein was identified only in the single isolate recovered from the PAR case.

Although we identified 51 OMPs using proteomic approaches that were also predicted by our previous bioinformatic analyses of the outer membrane proteomes of avian isolate Pm70 and porcine isolate 3480, there remained 47 putative predicted OMPs that were not identified in the present study. These included 22 lipoproteins, 10 transmembrane β-barrel proteins, four transmembrane β-barrel lipoproteins and 11 unknown proteins. Eight of these proteins (hemin receptor PM1282, haemoglobin receptors PfhR and HmbR, uncharacterized TonB-dependent receptors PM0745 and PM1081, OmpH_3, outer membrane efflux lipoprotein IbeB and polysaccharide export protein Wza) function in transport, four proteins (NlpD-like protein, NlpC, peptidase M48B family protein and a sialidase NanB) have enzymatic activity, three proteins (outer membrane lipoprotein-insertion protein LolB, membrane-bound lytic murein transglycosylases MltB and MltC) are involved in outer membrane biogenesis, three proteins (competence protein and secretin ComE, autotransporter adhesins Hsf_1 and Hsf_2) are involved in adherence, and one protein (Mod_2) has DNA binding activity. Twenty-eight of the 47 unidentified proteins have unknown functions and these include 16 lipoproteins, two transmembrane β-barrel proteins, four transmembrane β-barrel lipoproteins and six unknown proteins.

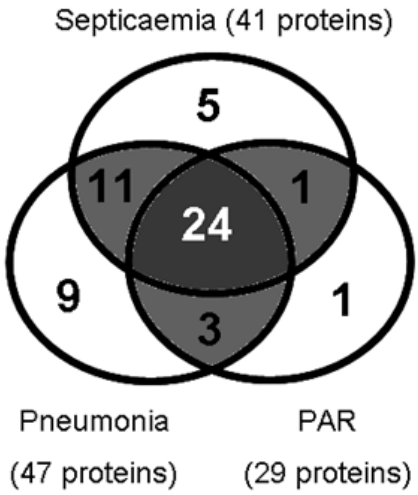


Figure 3-14. Association of OMPs of *P. multocida* isolates with three disease syndromes: septicaemia associated with two avian isolates, pneumonia associated with two bovine, one porcine and two ovine isolates, and PAR associated with one porcine isolate. Numbers indicate the number of proteins. The total numbers of OMPs are shown in parentheses.

The present study identified 27 putative contaminant proteins (33% of the identified proteins) and these included six cytoplasmic proteins (PM0979, NqrA, NqrF, Ef-Tu-A, PM1175 and LosA), 11 inner membrane proteins (PntA, YajC, CydA, PtnC, NqrC, HexC, PM1299, PM1683, PM1132, PM0876 and PM1918), three periplasmic proteins (Cah, 21 kDa hemolysin precursor and TolB), and seven proteins of unknown localization (TadG, amino acid adenylation protein, hypothetical protein CGSHi22421_00657, cell division protein FtsK, type II secretion protein Exig_0881 homologue in *Exiguobacterium sibiricum*, uncharacterized protein PM0132 and lipoprotein PM0553). The gel-free method identified a total of 25 contaminant proteins and seven proteins per isolate on average, whereas the gel-based method identified only 12 contaminant proteins in total and three proteins per isolate on average.

3.3.5 Comparative proteomic analyses of the cell envelope protein profiles

The analysis of cell envelope profiles was studied to identify proteins lost by Sarkosyl extraction. Cell envelopes of the eight isolates of *P. multocida* were prepared by ultracentrifugation of ultrasonicated cells. These contain the inner membrane-periplasmic-outer membrane complex. The cell envelope proteins were separated by two options of 1D SDS-PAGE: running until the samples reached the bottom of the gel (**Figure 3-15A**) and running until the samples migrated into the resolving gel for only 1-1.5 cm (**Figure 3-15B**). For the latter option, all of the bands (containing all of the proteins) were cut out and divided into equal gel slices before in-gel trypsin digestion and LC-MS/MS analysis. The cell envelope profiles were more complex than the OMP profiles. However, the highly abundant major OMPs, OmpA and OmpH, can be clearly seen on this gel (**Figure 3-15A**).

3.3.5.1 Comparison of cell envelope proteins associated with different host species

The extracted cell envelopes of the eight isolates of *P. multocida* contained cytoplasmic proteins, cytoplasmic membrane proteins, periplasmic proteins, outer membrane proteins, extracellular protein and unknown proteins (**Figure 3-16 and Appendix Table 3-6**). An average of 223 proteins was identified in all eight isolates. The majority of these proteins were cytoplasmic proteins (136 proteins on average), while the average numbers of cytoplasmic and outer membrane proteins were comparable (37 cytoplasmic and 31 outer membrane proteins) (**Figure 3-16**). A small proportion of periplasmic (six proteins on average) and extracellular (one protein on average) proteins were also identified. Fifty-one OMPs were identified across all eight isolates, accounting for 49% of the predicted outer membrane proteome of the avian strain and 45% of the predicted outer membrane proteome of the porcine strain (**Figure 3-17**).

Comparison of the identified OMPs obtained by the extraction of the cell envelope across four animal hosts showed that 27 proteins were found in isolates associated with all four animal hosts (**Figure 3-18**). These included FadL, Lpp/Pcp, LppC, NlpB, Oma87, OmpW, MetQ/PlpA, Pal, OmpH_1, OmpH_2, OmpA, ComL, Plp4, RlpB, TolC, HexD, HlpB, MipA/OmpV, YccT, TadD, EstA, Skp and hypothetical proteins PM0016, PM1886, PM0442, PM1323 and PM1827. Six proteins were identified in isolates from three animal hosts: LppB/NlpD, hypothetical proteins PM1826 and PM1798 in bovine, porcine and ovine isolates; RcpA and hypothetical protein PM0674 in avian, bovine and ovine isolates; and RcpC in avian, porcine and ovine isolates. Seven proteins were identified in isolates from two animal hosts: Lipoprotein E/OmpP4 in bovine and porcine isolates; hypothetical proteins PM1077 and PM1543 in avian and ovine isolates; OmpLA and HgbA in porcine and ovine isolates; TbpA and Slp in bovine and ovine isolates. Lastly, 11 proteins were

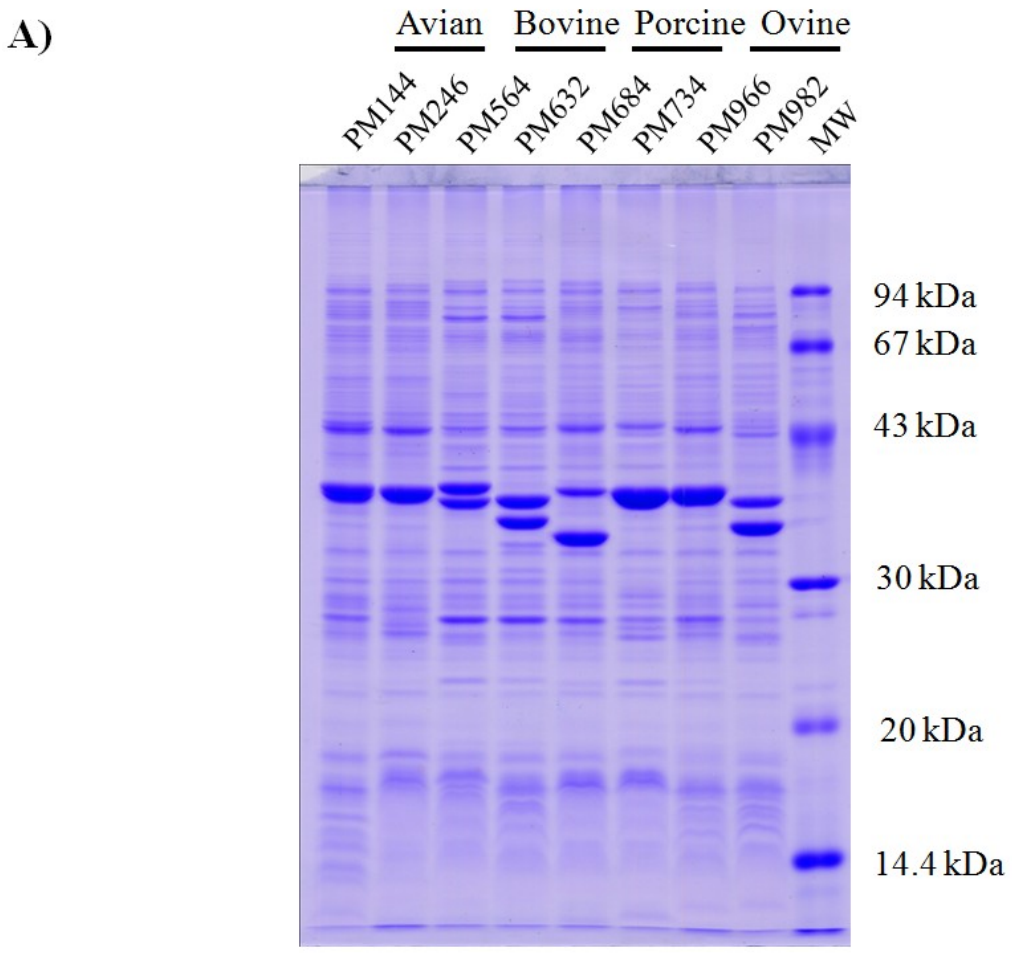


Figure 3-15. The cell envelope profiles of the eight isolates of *P. multocida* (A). The same samples were run into the resolving gel for 1-1.5 cm and all of the bands were cut and subjected to in-gel trypsin digestion before the LC-MS/MS analysis (B).

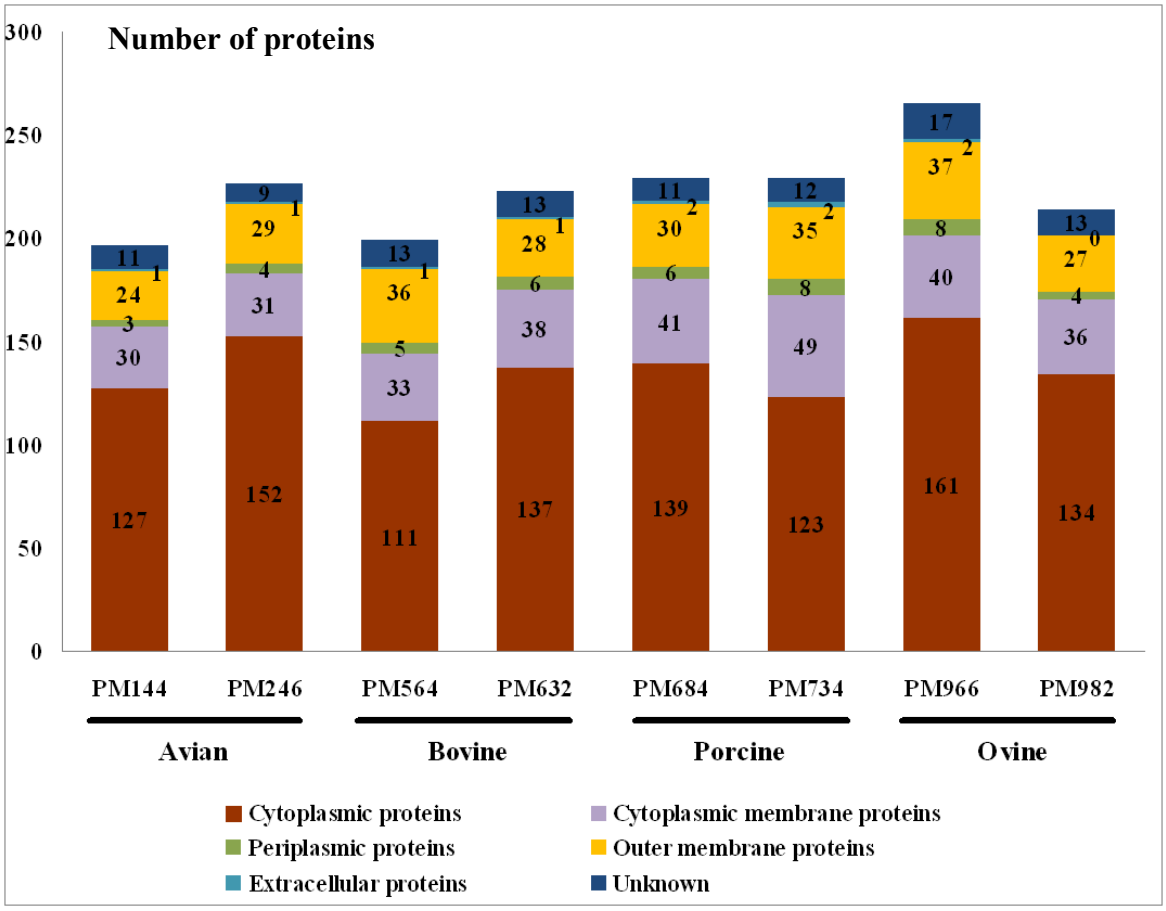


Figure 3-16. Comparison of the numbers of proteins including cytoplasmic, inner membrane, periplasmic, outer membrane and extracellular proteins, identified in the cell envelope profiles of eight isolates of *P. multocida*. The numbers in the bar graph indicate the number of proteins.

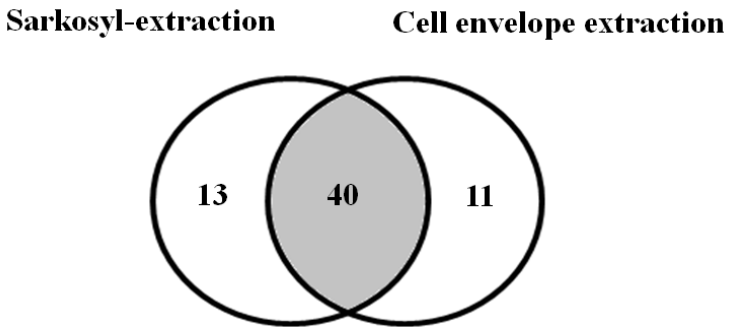


Figure 3-17. Comparison of the total number of different OMPs identified in all isolates of *P. multocida* using Sarkosyl and cell envelope extraction. The area shaded in grey represents the number of proteins identified by both methods and the area without shading represents the number of proteins identified by only one method.

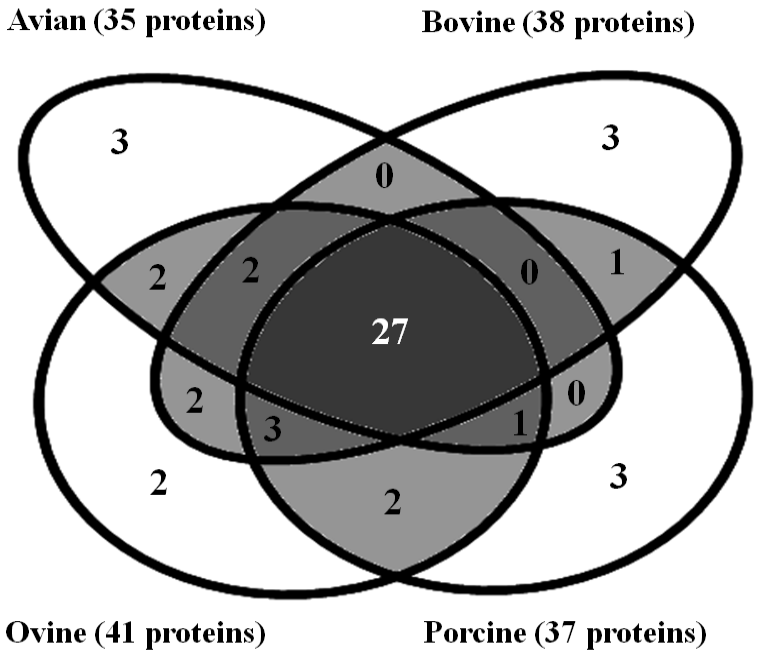


Figure 3-18. Comparison of the OMPs identified from the cell envelope profiles in isolates from four different animal hosts (avian, bovine, porcine and ovine). The OMPs identified in four, three, two and one animal hosts are shown in dark gray-, medium dark gray-, light grey-shaded and non-shaded areas, respectively. Numbers indicate the number of proteins. The total number of OMPs associated with each host is shown in parentheses.

identified in isolates from one animal host: Opa, RcpB and a hypothetical protein PARMER_00516 in the avian isolates; NlpC, hypothetical protein PM1578 and PM0663 in the bovine isolates; MltC and hypothetical proteins PM1809 and PM1428 in the porcine isolates; and VacJ and LolB in the ovine isolates.

3.2.5.2 Comparison of the identified cell envelope proteins with the identified OMPs

Comparison of the OMPs identified by Sarkosyl extraction to those identified by the extraction of the cell envelope showed that 40 OMPs were detected by both methods in at least one of the eight isolates. Thirteen OMPs were identified in at least one isolate by only the Sarkosyl extraction. These included OstA, GlpQ, a conserved hypothetical protein PM1515, LspB_1, LspB_2, Mce/PqiB, HasR, TonB-dependent receptor PM0803, two HgbB proteins, HemR, YtfN-like protein, an autotransporter PM1717. Twelve of these proteins were transmembrane β-barrel OMPs. Eleven proteins were detected only by the extraction of the cell envelope in at least one isolate. These included HlpB, Skp, MltC, LolB, NlpC and hypothetical proteins PM1323, PM0674, PM1798, PM1578, PM1543 and PARMER_00516. MltC and LolB involved in biogenesis and integrity of outer membrane. NlpC is a cell wall peptidase. Functions of other proteins are unknown. Eight of these proteins were lipoproteins. Combining the OMPs identified by these two methods yielded 64 OMPs (**Figure 3-17**). These account for 62% of the predicted avian strain outer membrane proteome and 57% of the porcine strain proteome. Detailed comparison of the number of OMPs identified by these two methods within the same isolate is illustrated in **Figure 3-19.**

Considering OMPs obtained from extraction of the cell envelope, 31 proteins were detected in some isolates by the Sarkosyl extraction and they were additionally identified in a few more isolates by the extraction of the cell envelope (**Figure 3-20**). The majority.

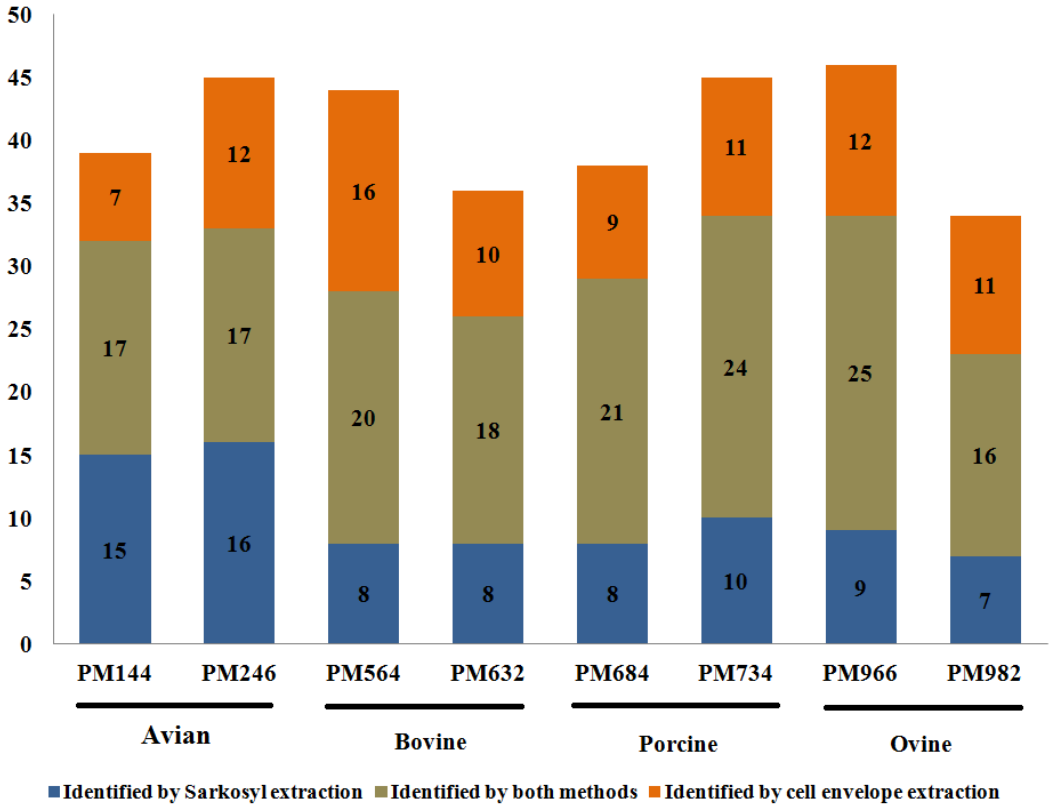


Figure 3-19. Comparison of the OMPs identified by the Sarkosyl and cell envelope extraction methods. The top part of each bar represents OMPs identified only by cell envelope extraction. The OMPs identified by both methods are shown in the middle part of each bar. The bottom part of each bar represents OMPs identified only by Sarkosyl extraction. Numbers indicate the number of OMPs detected in each part.

(22) of these proteins were lipoproteins (**Figure 3-20**). Eight proteins were identified in five, six or seven isolates from the Sarkosyl extracts, and analysis of the cell envelope helped the identification of these proteins in a few more isolates. Thirteen proteins were previously identified in one, two and three isolates, and the analysis of the cell envelope increased the identification of these proteins in a few or more isolates. Ten OMPs were additionally identified from only the cell envelope analysis (**Figure 3-20**).

Comparison of the 64 OMPs (**Table 3-5**) identified by Sarkosyl extraction and the extraction of the cell envelopes across all four animal hosts revealed that 36 OMPs were common in all animal hosts (FadL, MetQ, Oma87, OmpA, OmpH_1, OmpH_2, Pal, TolC, Lpp/Pcp, OmpW, HexD, NlpB, ComL, EstA, MipA/OmpV, Lipoprotein E/OmpP4, LppB/NlpD, SmpA, HgbA, Plp4, LppC, RlpB, TadD, VacJ, LptD/Imp/OstA, YccT, GlpQ, HlpB, OmpLA,Skp, hypothetical proteins PM0016, PM0442, PM1826, PM1827, PM1809 and PM1323); eight OMPs were associated with three animal hosts (TbpA, RcpA, RcpC, LspB_1, Mce/PqiB, hypothetical proteins PM0674, PM1515, PM1798,); nine OMPs with two animal hosts (Slp, HasR, LspB_2, two HgbB proteins and hypothetical proteins PM1077, PM1428, PM1543, PM0803); and 11 OMPs from only animal host (**Figure 3-21**). Five OMPs were detected only in the avian isolates: Opa, RcpB, hypothetical protein PARMER_00516, OmpL41/YtfN-like protein and an autotransporter PM1717. Three OMPs (NanH, NlpC and hypothetical protein PM1578) were identified only in the bovine isolates, whereas two proteins (MltC and HemR) were identified only in the porcine isolates and one protein (LolB) was identified only in the ovine isolates.

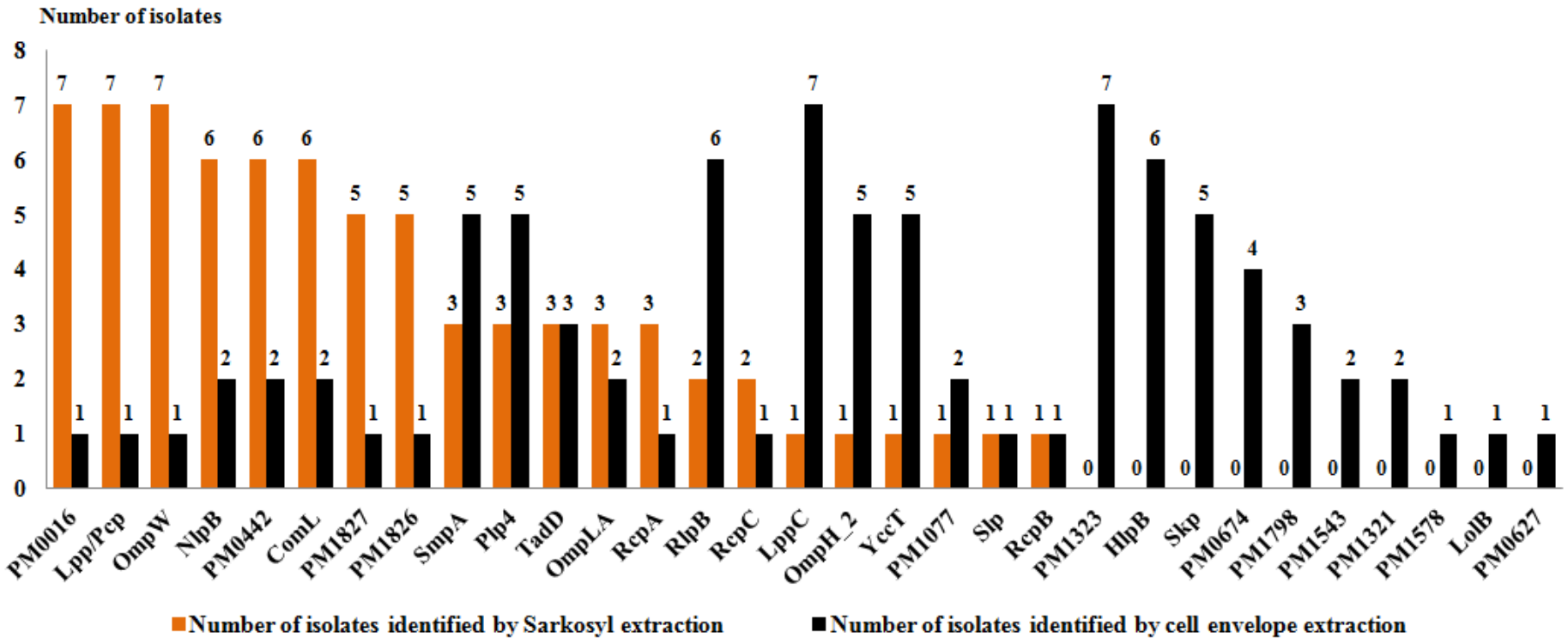


Figure 3-20 Comparison of the number of isolates from which OMPs were identified by the Sarkosyl and/or cell envelope extraction methods. Certain OMPs were identified in a few isolates by Sarkosyl extraction, but extraction of the cell envelope additionally detected them in a few or more isolates. New OMPs were also identified by the analysis of the cell envelope profiles.

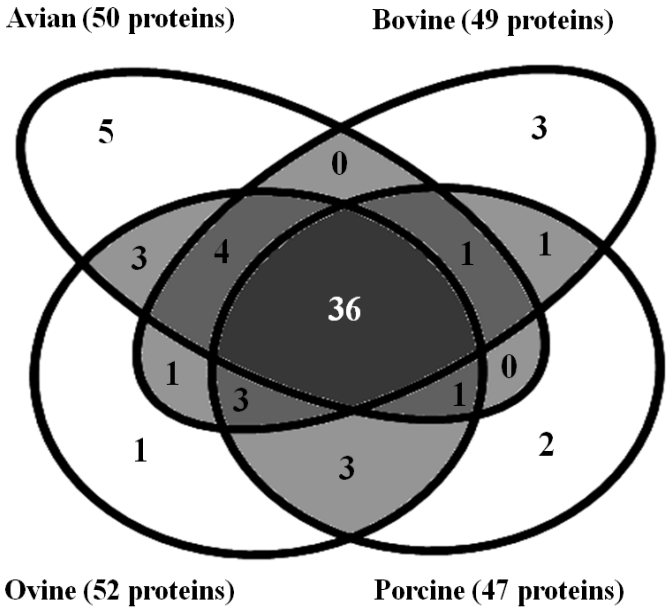


Figure 3-21. A comparison of the combined OMPs identified from both Sarkosyl and cell envelope extraction methods in isolates from four different animal hosts (avian, bovine, porcine and ovine). The OMPs identified in four, three, two and one animal hosts are shown in dark gray-, medium dark gray-, light gray-shaded and non-shaded areas, respectively. Numbers indicate the number of proteins. The total number of OMPs associated with each host is shown in parentheses.

Table 3-5. Comparison of the OMPs identified from the Sarkosyl and cell envelope extraction methods of the eight isolates of *P. multocida.*

No	ID	Name	Different isolates of *P. multocida*															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction
1	PM1069	FadL	+	+[1]	+	+[1]	+	+	+	+[1]	+	+[1]	+	+	+	+	+	+
2	PM1730	PlpA/MetQ	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3	PM1992	Omp87/Oma87	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
4	PM0786	OmpA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
5	PM0388	OmpH_1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
6	PM0966	Pal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
7	PM0016	Lipoprotein PM0016	+	+	+	+[2]	+	+	+	+[2]	+	+	+	+[2]	+	+	+	-
8	PM0527	TolC	-	+	+	+	+	+	+	+	+	+[1]	+	+	+	+[2]	+	+

'+' = positive identification for cell envelope and identified by gel-based & gel-free for the OM fraction; '+[1]' = identified by gel-based; '+[2]' = identified by gel-free; '-' = no identification; '*'=identified only in OM fraction; '**'= identified only in cell envelope fraction

Table 3-5. (continued)

No	ID	Name	Different isolates of *P. multocida*															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction
9	PM0554	Lpp/Pcp	+	+[1]	+	-	+	+[1]	+	+[1]	+	+[1]	+	+[1]	+	+[1]	+	+[1]
10	PM0331	OmpW	+	+[1]	+	-	+	+[1]	+	+[1]	+	+[1]	+	+[1]	+	+	+	+[1]
11	PM0778	HexD	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+
12	PM1050	NlpB	+	+[2]	+	-	+	-	+	+[2]	+	+[2]	+	+[2]	+	+[2]	+	+[2]
13	PM0442	Lipoprotein PM0442	-	+[2]	+	+[2]	+	+[2]	+	+[2]	+	-	+	+[2]	+	+[2]	+	-
14	PM1720	ComL	-	+[2]	+	+[1]	+	+[2]	+	-	+	-	+	+[2]	+	+[2]	+	+[2]
15	PM0076	EstA	-	+[1]	+	+	+	+[1]	-	+[2]	+	+	-	+[1]	+	+	+	+
16	PM0998	MipA/OmpV	+	+	-	-	+	+	+	+	+	+[2]	+	+	+	+	-	+[1]
17	PM1064	Lipoprotein E/OmpP4	-	+[2]	-	+[2]	+	+	+	+[2]	+	+	+	+	-	+[2]	-	+[2]
18	PM1614	LppB/NlpD	-	+	-	+	+	+[1]	-	-	+	+	+	+	+	+[1]	+	+[1]

Table 3-5. (continued)

No	ID	Name	Different isolates of *P. multocida*															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction
19	PM1886	SmpA	+	+[1]	+	-	+	-	+	-	+	+[1]	+	+[1]	+	-	+	-
20	PM0300	HgbA	-	+	-	+[1]	-	+[1]	-	+	+	+	+	+	+	+	-	+
21	PM1827	Hypothetical protein PM1827	+	+[2]	+	+[2]	+	-	-	-	-	+[2]	+	+[2]	+	+[2]	-	-
22	PM0586	Plp4	-	+[2]	+	+	+	-	+	-	+	-	+	-	+	+[2]	+	-
23	PM1826	Hypothetical protein PM1826	-	+[2]	-	+[2]	+	+[2]	+	-	-	-	+	+[2]	+	+[2]	-	-
24	PM0646	LppC	+	-	+	-	+	+[1]	+	-	+	-	+	-	+	-	+	-
25	PM1809	Omp85 family PM1809	-	+	-	+	-	+	-	-	+	+	+	+	-	+	-	+
26	PM1215	RlpB	+	-	-	+[2]	+	-	+	-	+	+[2]	+	-	+	-	+	-
27	PM0846	TadD	+	+	+	+	+	-	+	-	-	-	+	-	+	+[2]	-	-
28	PM1501	VacJ	-	+[2]	-	+	-	+[2]	-	+[2]	-	+[2]	-	+[2]	+	+[2]	-	+[2]

Table 3-5. (continued)

No	ID	Name	Different isolates of *P. multocida*															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction
29*	PM1600	LptD/Imp/OstA	-	+[1]	-	+	-	+[1]	-	+[2]	-	+[1]	-	+	-	+	-	+[1]
30**	PM1323	Lipoprotein PM1323	+	-	+	-	+	-	-	-	+	-	+	-	+	-	+	-
31	PM0389	OmpH_2	+	-	+	-	+	-	-	-	+	+[1]	+	-	+	-	-	-
32	-	TbpA	-	-	-	-	+	+	+	+	-	+[2]	-	-	-	-	+	+
33	PM0395	YccT	-	-	+	-	+	-	+	+[2]	+	-	-	-	+	-	+	-
34*	PM1444	GlpQ	-	+	-	+	-	+[2]	-	+	-	+[2]	-	+[2]	-	+	-	-
35**	PM1805	HlpB (PM1805)	+	-	-	-	+	-	+	-	+	-	+	-	+	-	-	-
36	PM1426	OmpLA	-	-	-	+[2]	-	-	-	+[2]	+	-	+	+[2]	+	-	-	-
37	PM0852	RcpA	-	+[2]	+	+	+	-	-	-	-	-	-	-	+	+[2]	-	-
38	PM0853	RcpC	-	-	+	+[2]	-	-	-	-	-	-	+	-	+	+[2]	-	-

Table 3-5. (continued)

No	ID	Name	Different isolates of *P. multocida*															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction
39**	PM1993	Skp/Omp p25	+	-	-	-	+	-	+	-	-	-	+	-	+	-	-	-
40**	PM0674	Lipoprotein PM0674	-	-	+	-	-	-	+	-	-	-	-	-	+	-	+	-
41	PM1025	Opa	+	+[1]	+	+[1]	-	-	-	-	-	-	-	-	-	-	-	-
42	PM1077	PM1077	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+	-
43*	PM1515	Conserved hypothetical protein PM1515	-	+[1]	-	-	-	+[1]	-	-	-	+[1]	-	+	-	-	-	-
44**	PM1798	Lipoprotein PM1798	-	-	-	-	+	-	-	-	-	-	+	-	+	-	-	-
45	PM1428	TonB-dependent receptor PM1428	-	-	-	-	-	-	-	-	-	-	+	+	-	+[2]	-	-
46	PM0708	Slp	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+	+[2]
47*	PM0056	LspB_1	-	-	-	+[1]	-	+	-	-	-	-	-	-	-	+[1]	-	-
48*	PM0266	Mce/PqiB	-	-	-	+[1]	-	+[1]	-	-	-	-	-	-	-	+[1]	-	-
49*	PM1622	HasR	-	-	-	-	-	-	-	+[1]	-	+[1]	-	+	-	-	-	-

Table 3-5. (continued)

No	ID	Name	Different isolates of *P. multocida*															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction
50**	PM1543	Hypothetical protein PM1543	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-
51	PM0663	NanH	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
52**	PM1321	MltC	-	-	-	-	-	-	-	-	+	-	+	-	-	-	-	-
53	PM0851	RcpB	+	-	-	+[2]	-	-	-	-	-	-	-	-	-	-	-	-
54*	PM0058	LspB_2	-	-	-	+[1]	-	-	-	-	-	-	-	-	-	+[1]	-	-
55*	PM0803	TonB-dependent receptor PM0803	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-	-
56*	PM0336	HgbB (PM0336)	-	-	-	-	-	-	-	-	-	-	-	+[1]	-	+[2]	-	-
57*	PM0337	HgbB (PM0337)	-	-	-	-	-	-	-	-	-	-	-	+[1]	-	-	-	+[1]
58*	PM0576	HemR	-	-	-	-	-	-	-	-	-	+[1]	-	+[1]	-	-	-	-
59**	PM1578	Lipoprotein PM1578	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
60**	-	TonB-dependent receptor PARMER_00516	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 3-5. (continued)

No	ID	Name	Different isolates of *P. multocida*															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction	Cell envelope	OM-fraction
61**	PM0246	LolB	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
62**	PM0627	NlpC/P60	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
63*	PM1808	OmpL41/YtfN-like protein	-	-	-	+[1]	-	-	-	-	-	-	-	-	-	-	-	-
64*	PM1717	OM autotransporter PM1717	-	-	-	+[1]	-	-	-	-	-	-	-	-	-	-	-	-
		Total number of proteins	24	32	29	33	36	28	28	26	30	29	35	34	37	34	27	23

3.4 Discussion

3.4.1 Comparison of different OMP preparation methods

Different concentrations of Sarkosyl were examined and 0.5% Sarkosyl was demonstrated to be the optimum concentration for the extraction of *P. multocida* OMPs. Higher concentrations of Sarkosyl seemed to solubilise certain OMPs such as OmpA, whereas lower concentrations were unable to sufficiently remove the inner membrane proteins, resulting in more complex OMP profiles. However, the optimum concentration of Sarkosyl could depend on bacterial species. For example, 2% Sarkosyl was used to extract OMPs of *Helicobacter pyroli* (Baik *et al.*, 2004) and 1% Sarkosyl was used to extract OMPs of *Moraxella catarrhalis* (Hays *et al.*, 2005). When Sarkosyl was compared with other detergents, Sarkosyl could cleanly extract the outer membrane more efficiently than other detergents used in the study. The explanation could be due to differences in the structures and properties of these detergents shown in **section 3.1.2.1.1**. The addition of EDTA to the Sarkosyl solution could disturb the stability of the outer membrane since it resulted in the loss of certain OMPs such as OmpA. Similarly, Sarkosyl extraction also outperformed spheroplasting method. The spheroplasting method used EDTA and lysozyme to destabilize the outer membrane and caused artificial releasing of outer membrane vesicles. EDTA chelates Mg^{2+} and Ca^{2+} ions from outer membrane and lysozyme digests peptidoglycan structure (Birdsell & Cota-Robles, 1967). Overall, these data provide further evidence to support previous findings (Ravaoarinoro *et al.*, 1994; Hobb *et al.*, 2009) that Sarkosyl extraction is the best outer membrane extraction method to prepare the cleanest *P. multocida* OMP fractions. However, there remains the possibility that certain outer membrane lipoproteins which are loosely associated with the outer membrane will be removed by Sarkosyl. This will be discussed in **section 3.4.3**.

3.4.2 Comparative proteomic analyses of the outer membrane proteome

The eight representative isolates of *P. multocida* recovered from cattle, sheep, pigs and poultry were grown in complete media and outer membrane-enriched fractions were prepared by Sarkosyl extraction. Sarkosyl selectively solubilizes the inner membrane and produces an insoluble fraction representing the outer membrane-peptidoglycan complex (Filip *et al.*, 1973). This method has previously been used, by ourselves and others, to prepare outer membrane fractions of *P. multocida* and other Gram-negative bacterial species and "clean" outer membrane-enriched fractions have been satisfactorily obtained (Kaur *et al.*, 2003; Davies *et al.*, 1992, 2003a, 2003b, 2003d, 2004; Chung *et al.*, 2007; Liu *et al.*, 2008; Hobb *et al.*, 2009). Davies *et al.* (2003a, 2003b, 2003d, 2004) have established that the OMP profiles of pairs of *P. multocida* isolates from the four hosts are highly stable after multiple subculture and after growth to the same stage of the growth cycle in BHIB. In particular, these authors used the stability and identity of OMP profiles of isolates representing the same clone to classify *P. multocida* into distinct OMP-types (Davies *et al.*, 2003a, 2003b, 2003d, 2004). In the present study, two further replicate OMP samples of the eight isolates were generated and these confirmed the stability and identity of profiles representing the same isolate; consequently, subsequent proteomic analyses were performed on one set of samples.

Thein *et al* compared five outer membrane preparation methods for *E. coli* strain BL21, which incorporated different combinations of lysis (lysozyme/EDTA and spheroblasting or French Press), membrane pelleting (ultracentrifugation or washing with chaotropic reagents), and inner/outer membrane separation (selective detergent treatment using Triton X-100 or sucrose density gradient centrifugation) (Thein *et al.*, 2010). Although these authors did not include Sarkosyl extraction in their study, their results suggested that use of a French Press followed by a carbonate wash was the most satisfactory outer membrane

extraction method with the least contamination (37% contamination with cytoplasmic proteins, mostly ribosomal proteins). Thein *et al* identified a total of 44 different OMPs using GelC-MS from the single *E. coli* strain and 34 OMPs on average in each method (Thein *et al.*, 2010). This figure is slightly higher than our average of 30 OMPs (54 proteins in total from eight isolates), but our method yielded fewer contaminant proteins (33% in total and 25.9% on average) and scarcely no ribosomal proteins. Hobb *et al.* evaluated nine different methods for the preparation of outer membrane-fractions of *Campylobacter jejuni* (Hobb *et al.*, 2009). These authors compared the use of different detergents and extraction buffers, the production of spheroplasts by lysozyme or sonication, and the use of sucrose density gradient centrifugation. The Sarkosyl extraction, gradient centrifugation and spheroplasting methods produced samples free of cytoplasmic and inner membrane proteins. However, these authors concluded that Sarkosyl extraction provided the purest outer membrane extracts and was the most reproducible method. Our Sarkosyl-extracted outer membrane fractions also contained low numbers of contaminant proteins (25.9% on average) and two separate OMP preparations gave almost identical OMP profiles. Thus, we demonstrated that Sarkosyl extraction is a simple, rapid method that provides a clean and reproducible outer membrane sample. However, the most effective outer membrane extraction method for one bacterial species may not be appropriate for another and complementary methods may increase the coverage of the outer membrane sub-proteome (Thein *et al.*, 2010).

Using a simple bioinformatic prediction framework, we have confidently identified 98 putative OMPs in the genome of the avian *P. multocida* strain Pm70 and 107 OMPs in the incomplete genome of porcine *P. multocida* strain 3480. In this Chapter, we analyzed the OMPs of eight *P. multocida* isolates using complementary proteomic methods and identified 54 putative OMPs. These represented 52% of the predicted avian outer membrane sub-proteome and 48% of the predicted porcine sub-proteome. Thirty-two

OMPs were identified by both gel-free and gel-based methods whereas 12 OMPs were identified by the gel-based and 10 by the gel-free methods alone **(Figure 3-11A)**. The gel-based approach identified a higher proportion of β-barrel proteins with a higher average molecular mass, whereas the gel-free approach identified a higher proportion of lipoproteins. Therefore, these complementary approaches resulted in better coverage of the outer membrane proteome, consistent with previously published work (Chung *et al.*, 2007; Cordwell *et al.*, 2008; Tan *et al.*, 2008; Hauberg *et al.*, 2010; Jungblut *et al.*, 2010; Kouyianou *et al.*, 2010; Thein *et al.*, 2010). Although 1-DE has lower resolution of protein separation than 2-DE, it has been suggested as a more appropriate method for the identification of membrane proteins (Boyce *et al.*, 2006; Chung *et al.*, 2007) because 2-DE has limitations for the separation of very hydrophobic proteins (e.g., inner membrane proteins). Very hydrophobic proteins can be unfolded and solubilized in 1-DE sample buffer which contains SDS (Rabilloud *et al.*, 2009) and multiple samples can be run and compared on the same 1-D gel. However, due to the lower hydrophobicity of OMPs (negative GRAVY score and a low number of α-helical strands) compared to inner membrane proteins, the separation of OMPs by 2-DE should not be problematic. Therefore, either 1-DE or 2-DE can be used to visualize the outer membrane proteome. Boyce *et al*. used two gel-based methods (1-DE followed by MALDI-TOF MS and 2-DE followed by LC-MS/MS) and identified 24 OMPs (24%) in avian *P. multocida* strain X-73 grown in complete medium (Boyce *et al.*, 2006). These authors identified 21 OMPs by 1-DE and 16 OMPs by 2-DE; 13 proteins were identified by both methods. In our study, we identified 21 of their 24 OMPs by 1-DE and two further OMPs (lipoprotein PM0442 and phospholipase OmpLA) by the gel-free method; only one OMP (lipoprotein PM1578) was not identified in the present study. These results confirm that 1-DE can be used for analyses of the outer membrane proteome. In the present study, we identified 54 OMPs using gel-based and gel-free methods and have improved the coverage of the *P. multocida* outer membrane sub-proteome, from 24% to 33-34% in the avian isolates and from 24% to

52% for all eight isolates, by combining data from the gel-based and gel-free methods, and by incorporating more diverse isolates associated with different animal hosts and diseases.

Comparison of the OMPs identified in isolates from the four animal hosts allowed us to distinguish between two major groups of proteins, namely core and host-associated proteins. Twenty-four core OMPs were identified and these were shared by isolates associated with all four animal hosts. These proteins have essential functions including outer membrane biogenesis and integrity, transport and receptor, and enzymatic activity as well as unknown functions. Presumably, the absence of these proteins will likely have deleterious effects on the bacteria and their presence is essential (Davie & Campagnari, 2009; Platz *et al.*, 2010). Six of the core proteins (OmpH_1, PlpA/MetQ, OmpA, Pal, Lipoprotein E/OmpP4 and PM0442) were predicted to be highly expressed according to their high CAI values (> 0.35). The high abundance of the two major OMPs, OmpA and OmpH_1, was clearly visible in our 1-D gels and this was consistent with their high CAI scores. The other four OMPs were lipoproteins which may have been partly solubilized by Sarkosyl extraction due to their loose association with the outer membrane. Certain of the core proteins exhibited molecular mass variation across the eight isolates. This suggests the probability of substantial nucleotide sequence diversity in the genes encoding these proteins and likely reflects complex evolutionary histories and/or variation in protein domain structure (Davies *et al.*, 2001; Davies & Lee, 2004a). It is interesting to speculate that such variation might reflect differing protein function in different animal hosts.

The eight representative isolates also possessed host-associated proteins that could be functionally categorized into three groups: adherence and colonization, TonB-dependent iron receptors and those of other diverse functions. Proteins involved in adherence and colonization included RcpA, RcpB, RcpC, TadD, LspB_1, LspB_2, Opa, NanH and an uncharacterized autotransporter PM1717. RcpA, RcpB, RcpC and TadD are encoded by genes of the *tad* locus which is part of a genomic island called the widespread colonization

island (WCI). These proteins are important for biofilm formation, colonization and pathogenesis in many bacteria and Archaea (Tomich *et al.*, 2007) but, noteably, were only expressed in the avian and ovine isolates. Although these proteins were predicted in the porcine genome, they were not identified in either of the porcine isolates studied. RcpA, RcpC and TadD may be expressed at low levels since their CAI scores are in the lowest group (0.199, 0.207 and 0.190, respectively). Similarly, LspB_1, LspB_2, NanH and PM1717 expression was not detected in either of the porcine isolates and NanH was not identified in either of the avian isolates although the genes encoding these proteins are present in both the avian and porcine strain genomes. Different levels of OMP expression in these representative isolates under identical growth conditions may be due to differences in gene regulation and/or post-translational mechanisms. Alternatively, these differences may be a consequence of strain-adaptation to different host species and to different niches within individual hosts. In support of this, Odenbreit *et al.* compared the expression of eight adhesin or adherence-associated genes in 200 clinical isolates of *Helicobacter pylori* (Odenbreit *et al.*, 2009). These authors showed that only two proteins were identified in all isolates, whereas the presence of other proteins was highly variable and essential for adaptation to an individual patient or niche. In the present study, the Opa protein, which is similar to the opacity protein of *Haemophilus influenzae* and belongs to the opacity porin family involved in host interaction (Dehio *et al.*, 1998), was identified only in the avian isolates. Although the Opa protein was predicted in the porcine strain genome, it was not identified in either of the porcine isolates in the present study. Interestingly, the avian and porcine *opa* genes are only 52% identical, suggesting that the Opa protein may have diverged as a result of adaptation to different host species. Dabo *et al.* suggested that extracellular matrix molecules on host cell surfaces act as common adherence sites for *P. multocida* isolates associated with different diseases and animal hosts (Dabo *et al.*, 2007). Different animal hosts and tissue types may trigger the expression of different adherence

molecules in host-specific isolates of *P. multocida* to facilitate attachment to, and colonization of, different host niches.

TonB-dependent iron receptors included HgbA (present in all isolates), HasR, HemR, two HgbB proteins, TbpA, PM0803 and PM1428 (detected in either the bovine, porcine or ovine isolates but not in the avian isolates). These proteins bind to different iron sources including haem, haemoglobin, haemoglobin-haptoglobin, and transferrin (Bosch *et al.*, 2004; Prado *et al.*, 2005b). Bosch *et al*. suggested that the possession of multiple iron receptors leads to increased levels of iron acquisition and prevents the lethal effects of mutations in these genes (Bosch *et al.*, 2004). In a previous study, PM0803 was shown to be selectively expressed in the avian strain X-73 under iron-limited and *in vivo* conditions (Boyce & Adler, 2006). However, in the present study, this protein was detected in bovine and porcine isolates grown in a complex growth medium. HasR, a member of a haemophore system, was also identified in bovine and porcine isolates, whereas two HgbB proteins and PM1428 were detected in porcine and ovine isolates. TbpA was identified in both bovine isolates and in a single ovine isolate by both gel-based and gel-free methods, and in one porcine isolate by the gel-free method alone. Noteably, TbpB was not identified in any of these isolates. In other bacterial species, *tbpA* is normally present with *tbpB* as part of the *tbpBA* operon. However, Ogunnariwo *et al*. demonstrated that bovine strains of *P. multocida* lack TbpB but possess a novel form of TbpA which is sufficient, by itself, for iron acquisition (Ogunnariwo & Schryvers, 2001). Shivachandra *et al*. (Shivachandra *et al.*, 2005) similarly detected *tbpA*, but not *tbpB*, in a porcine strain, whereas Ewers *et al*. (Ewers *et al.*, 2006) were unable to detect either gene among a selection of porcine strains. Thus, we have not only confirmed that bovine and certain porcine isolates possess TbpA but not TbpB, but have extended this observation by demonstrating that certain ovine isolates similarly possess TbpA alone. It would seem likely that this "novel" TbpA protein may have been acquired independently by certain

ovine and porcine strains, from bovine strains, by horizontal gene transfer. HemR was identified in the porcine isolates but not in any of the other isolates suggesting that this iron-uptake mechanism is specific for porcine isolates. The expression of different iron-uptake receptor proteins among these eight isolates might reflect differences in their abilities to adapt to new host environments containing various types and quantities of iron-containing compounds. Evidence for this is provided by Klitgaard *et al.* who compared gene expression in six serotypes of *Actinobacillus plueropneumoniae* in response to iron limitation (Klitgaard *et al.*, 2010). In addition to their common set of iron-regulated proteins, the authors demonstrated that the expression of three putative haemoglobin-haptoglobin binding proteins was lowest in the least virulent serotype. These results suggested that increased expression of these proteins may assist bacterial adaptation and disease pathogenesis.

The third group of host-associated proteins, those having various diverse and unknown functions, included phospholipase A/OmpLA, Mce/PqiB, SmpA, RlpB, Plp4, Slp, HlpB, OmpL41/YtfN-like protein, YccT, LppC, OmpH_2, and hypothetical proteins, PM1827, PM1515 and PM1134 (**Tables 3-3 and 3-4**). Phospholipase A has an enzymatic role in maintaining asymmetry of the outer membrane. Only two of the host-associated proteins represented in this third group (i.e. phospholipase A and OmpH_2) were previously identified by Boyce *et al*. (Boyce *et al.*, 2006). Certain proteins were functionally uncharacterized or have homologues of unknown functions in other bacterial species. A small number of proteins, including SmpA and RlpB, are involved in outer membrane biogenesis and integrity and should probably have been identified in all isolates (but were only detected in some isolates). There are three possible reasons to explain this finding. First, SmpA (15.5 kDa) and RlpB (18.9 kDa) are small lipoproteins which may be loosely associated with the outer membrane and could be lost during Sarkosyl extraction. Second, the CAI values of these two proteins are not high (0.234 for SmpA and 0.225 for RlpB)

suggesting that the proteins are of low abundance in the outer membrane. Third, these proteins simply may not be expressed under the prevailing growth conditions.

Twenty-seven of the proteins identified in the present study (representing 33% of the total) were putative contaminant proteins; these included cytoplasmic proteins, inner membrane proteins and periplasmic proteins. There are a number of reasons to explain this observation. Certain protein complexes or pathways involve interactions of proteins from different localizations which probably remained associated with the OMPs during Sarkosyl extraction. Such proteins include TadG, TolB, HexC and AcrB. Seven proteins (PM0979, YajC, Ef-Tu-A, NqrC, NqrF, 21 kDa hemolysin and PM1175) had CAI values greater than 0.29, suggesting that these proteins were highly abundant and may not be completely removed during the extraction step. Alternatively, some of these proteins could occupy multiple localizations.

In this study, outer membrane protein identifications from either the gel-based or gel-free methods were based on significant protein scores from the MASCOT database with at least one peptide hit. Many proteomics studies have accepted reliable protein identification when two or more peptides were identified. However, Gupta and Pevzner (2008) examined this two-peptide rule on protein identification by using multiple search tools and data sets. The authors found that this rule increased false discovery rates when compared to the single-peptide rule and the number of proteins identified was reduced. An estimation of error rates was suggested as an alternative to the two-peptide rule. Reproducibility of proteomic methods is another important issue for the quality of the conclusions drawn from the data. For the gel-based methods, the present study detected the same OMPs (mostly the core OMPs) from gel bands excised from various formats of 1D SDS gels. Occasionally, the same OMPs were identified from multiple bands within the same isolate of the same gels. The reason could be due to fragmentation and the large amount of these proteins in the samples. The overlapping of the OMPs detected from

different gels or bands indicated a certain degree of reproducibility. For the gel-free methods, the use of in-solution trypsin digestion can yield results with a low degree of reproducibility because the peptide production depends on accessibility of trypsin to the protein samples in the solution. Utilizing LC-MS/MS may introduce many possible sources of variability such as complexity of the analyses, variation in retention times of the liquid chromatography column, different mass spectrometric instruments, sample complexity and contamination (Tabb *et al*., 2010). Therefore, protein identification obtained by the gel-based methods could be more reproducible than those identified by the gel-free methods together with the use of LC-MS/MS. Analyses of the same sample multiple times or by using different proteomic methods (the gel-based and gel-free methods) will increase the degree of reproducibility of the protein identification.

Significantly, 47 putative predicted OMPs from the avian strain genome and 56 putative predicted OMPs from the porcine strain genome were not detected in the proteomic analysis. The explanations for this are similar to above. Oldfield *et al.* used Triton X-100 to prepare an outer membrane-fraction of *A. pleuropneumoniae* (Oldfield *et al.*, 2008). These authors found that ComL, LolB and LppC were predominant in their outer membrane-fractions. We identified ComL in six isolates and LppC in one isolate but did not identify LolB. Whatever the reasons, our findings clearly emphasize that no single method, including Sarkosyl extraction, can provide a complete outer membrane sample. Thus, complementary preparation methods are likely to be necessary for the optimum identification of OMPs. However, some of the proteins that were not detected might be produced in greater abundance under different growth conditions or be induced by specific stimuli (Boyce *et al.*, 2004; Papasotiriou *et al.*, 2008). The trimeric autotransporter Hsf was detected in *P. multocida* at transcriptomic levels under nutrient-limited and *in vivo* conditions (Paustian *et al.*, 2002; Baltes & Gerlach, 2004). However, we did not identify this protein after growth in complex medium. Some OMPs may be expressed during iron-

limited and *in vivo* growth conditions (Paustian *et al.*, 2001; Baltes & Gerlach, 2004; Boyce & Adler, 2006; Boyce *et al.*, 2006). The expression of some virulence genes (*wza*, *hexD*, *hgbA*, *pm1428*, *tadD*, and *lspB*) of a bovine isolate of *P. multocida* was affected by the presence of different antibiotics (Melnikow *et al.*, 2008). Thus, in addition to the method used to obtain the outer membrane sample, consideration must also be given to the growth conditions of the bacterium to allow identification of additional OMPs.

3.4.3 Comparative proteomic analyses of the cell envelope proteome

Using Sarkosyl to solubilise the inner membrane from outer membrane could result in the possible loss of certain loosely-associated outer membrane lipoproteins. The aim of studying the cell envelope profiles was to determine if any OMPs were lost by Sarkosyl extraction. The cell envelope profiles of the eight isolates of *P. multocida* contained mostly cytoplasmic proteins with comparable amount of inner and outer membrane proteins, and a few periplasmic and extracellular proteins. The cell envelope profiles were prepared by ultracentrifugation of the sonicated fractions which was one step prior to Sarkosyl extraction. Identification of OMPs from the cell envelope profiles yielded 51 OMPs in total, representing 57% to 62% of the predicted outer membrane proteomes from avian and porcine genomes, respectively. Compared with the OMP profiles prepared by Sarkosyl extraction, 40 proteins were identified by both methods, 13 proteins by Sarkosyl extraction, and 11 proteins by extraction of the cell envelopes. The cell envelope extraction method allowed the identification of additional outer membrane lipoproteins which were possibly solubilised by Sarkosyl because they are loosely associated with the outer membrane. On the other hand, 12 transmembrane β-barrel OMPs were lost in the extraction of the cell envelope, whereas these proteins were enriched during Sarkosyl extraction. The reason could be due to the low abundance of certain of these proteins compared to the highly abundant inner membrane and cytoplasmic proteins. Complexity

of the cell envelope samples was much higher than the Sarkosyl-extracted fractions. If the resolution and sensitivity of the mass spectrometer is not high enough, there is the probability of losing some peptides during the mass spectrometric analysis. Although Sarkosyl extraction produced reasonably clean outer membrane samples, it could be biased towards the transmembrane proteins. Therefore, combining the OMPs identified by these two methods could overcome this bias and increase the confidence of the OMP identification overall. This combination also improves the chance of identifying new OMPs. Ten new OMPs were identified by extraction of cell envelopes.

The comparison of the OMPs identified from these two methods across all four animal hosts resulted in changing of the number of proteins on some intersected area of the Euler's diagram in **Figure 3-21**. The number of core OMPs increased from 24 proteins to 36 proteins. These additional proteins were SmpA, Plp4, LppC, RlpB, TadD, OmpH_2, YccT, HlpB, OmpLA, Skp, Hypothetical proteins PM1827 and PM1323. However, the expression of certain OMPs from both methods remained the same. This suggested that these proteins are possibly host-specific or isolate-specific OMPs. An adherence protein, Opa was identified only in the avian isolates by both Sarkosyl extraction and extraction of cell envelopes. The other protein, sialidase (NanH) was identified only in bovine isolate PM632 by these two methods. Some of the host-restricted OMPs were first identified in specific hosts or isolates from Sarkosyl extraction and these proteins were further identified in additional hosts or isolates after extraction of cell envelopes. For example, TadD and RcpA which are involved in biosynthesis of Flp pili were identified in avian and ovine isolates by Sarkosyl extraction. However, these proteins were additionally identified in bovine and porcine isolates after extraction of cell envelopes. The avian and ovine isolates have non-mucoid colonies, whereas bovine and porcine isolates have watery mucoid colonies. These adherence proteins (Opa, TadD and RcpA) are frequently associated with non-mucoid isolates. However, the effect of capsule and extracellular

polysaccharide on Sarkosyl extraction has never been examined. These data suggest the possible loss of these proteins in some isolates during Sarkosyl extraction. Thus, the combination of these two extraction methods could confidently increase the number of identified OMPs.

In summary, the outer membrane fractions of eight *P. multocida* isolates associated with different animal hosts and disease syndromes grown in a complex growth medium were prepared by Sarkosyl extraction and complemented with extraction of cell envelopes. Complementary gel-based and gel-free proteomic methods were applied in combination with mass spectrometric techniques and enabled the identification of 54 different OMPs representing 52% of the predicted avian outer membrane sub-proteome and 48% of the predicted porcine sub-proteome from Sarkosyl extraction. Combining Sarkosyl extraction with extraction of cell envelopes identified 64 OMPs overall. Comparative outer membrane proteomics of these eight isolates identified a group of 36 core proteins and various host-specific proteins. Further studies of these proteins may shed light on how these bacteria adapt to different host species and niches, and cause different types of infections.

Chapter 4: Comparative outer membrane proteomic analyses of *P. multocida* isolates grown under different growth conditions

4.1 Introduction

The previous chapter compared the outer membrane proteomes of the eight representative isolates of *P. multocida* grown under *in vitro* growth condition in a complex growth medium (brain heart infusion broth). This chapter will focus on the changes to the outer membrane proteomes of these isolates in response to different *in vivo*-like conditions. These conditions include different stages of growth, growth with different rates of aeration, growth under iron limitation, growth in serum and growth under biofilm conditions.

4.1.1 Introduction to in vivo-like growth conditions

The response of bacteria grown in a complex growth medium will differ enormously from growing them *in vivo*. Poobalane *et al.* (2008) found different protein expressions while culturing *Aeromonas hydrophila* in a complete medium compared to *in vivo*. However, conducting *in vivo* experiments is a costly multi-step process and insufficient bacterial cells are usually obtained. *In vivo*-like growth conditions can be reproduced by at least partially, manipulating the *in vitro* conditions. Bacteria have to encounter many stress factors in order to grow within host environments. Nutrients essential for bacterial growth are limited *in vivo*, but are fully supplied *in vitro* in complex medium (Lorian, 1989). Deprivation of certain nutrients (e.g. iron) *in vitro* can mimic the *in vivo* growth conditions. Bacteria will be killed by the host defensive mechanisms (e.g. complement, antibodies and white blood cells) within the host blood stream or body fluids, whereas none of these mechanisms are found *in vitro* in the complex medium. During the *in vivo* infection process, bacteria have to be able to colonize and adhere to host surfaces. These surfaces provide a support for biofilm formation and multiplication of bacteria that resemble the

support provided by an agar surface *in vitro* (Lorian, 1989). Identifying the differences between growth in complex media compared to *in vivo*-like *in vitro* growth conditions will provide a better understanding of disease pathogenesis and host interactions (Poobalane *et al.*, 2008). Virulence proteins may be expressed when bacteria are grown under *in vivo*-like conditions compared to their culture in a complex medium. Ebanks *et al.* (2004) demonstrated similar patterns of OMP expression while culturing *A. salmonicida* under iron-restricted and *in vivo* growth conditions. Certain proteins were absent when *A. salmonicida* was grown in a complete iron-replete medium. Melnikow *et al.* (2005) identified differentially expressed genes in *Haemophilus parasuis* grown under different *in vivo*-like *in vitro* conditions including iron limitation, acidic and temperature stress, and microaerobic conditions. These authors reported genes that were up-regulated under all stress conditions and those which were specific to one or more conditions, suggesting that these genes were potential virulence and host-adapted factors in *H. parasuis*.

4.1.2 Response of P. multocida to different growth conditions

The response of *P. multocida* to different growth conditions has been studied at transcriptomic and proteomic levels. This response may be dependent on strains and animal host niches (Boyce & Adler, 2006). This was supported by Diallo & Frost (2000) who examined the survival of 35 avian strains of *P. multocida* in chicken serum and found that 27 strains were resistant to serum killing (some of these grew rapidly) whereas eight strains were susceptible. Similarly, the same strain of *P. multocida* showed variable levels of resistance to sera from different animals (Muhairwa *et al.*, 2002). At transcriptomic level, Paustian *et al.* (2001) compared levels of *P. multocida* gene expression during growth under iron-replete and iron-limited conditions and showed that certain OMPs altered their levels of expression e.g. HemR, HgbA (PM0300), HgbB (PM0336), haemin binding receptor PM1078, PM0803, haemoglobin receptor PM0741, ComD (PM1226) and OmpW. Paustian *et al.* (2002) compared levels of *P. multocida* gene expression during

growth under nutrient-rich and nutrient-limited conditions. These authors showed that 669 genes changed their levels of expression. These included genes encoding certain OMPs e.g. Hsf, OmpW, lipoproteins PM1518 and PM1926, and hypothetical proteins PM0674, PM1886 and PM1936. At the proteomic level, Boyce *et al.* (2006) compared the OMP expressions of an avian strain of *P. multocida* grown under *in vitro* (iron-replete and iron-limited) and *in vivo* (within bloodstream of infected chicken) growth conditions. These authors identified an OMP (PM0803) which was up-regulated during growth *in vivo* and under iron-limited conditions; the protein was not identified during growth under iron-replete conditions.

4.1.3 Effects of different growth conditions on the expression of OMPs

Bacteria need to respond when their growth conditions (host or external environments) change to maintain homeostasis. Detection of environmental changes can be observed by altered OMP expression. This section will review the effects of different growth conditions on the expression of OMPs.

4.1.3.1 Different growth stages

Bacterial growth in batch culture can be divided into four stages: lag phase, logarithmic phase, stationary phase and death phase (Cowan, 2012). Bacteria adapt themselves to their new environment during the lag phase. Once adapted, they rapidly grow at an exponential rate during the logarithmic phase which continues until there is a lack of adequate nutrients. During stationary phase, bacterial growth enters an equilibrium stage where the rate of cell division is equivalent to the rate of cell death. After that, more bacterial cells begin to die and the number of living cells declines during the death phase due to depleted nutrients and oxygen, and accumulated wastes.

Certain OMPs are differently expressed during different stages of the growth (Gallot-lavalme *et al.*, 1995). Evans and Poole (1999) showed that expression of an outer membrane efflux protein OprM (a homologue of TolC in *P. multocida*) was low during the lag phase and increased during the logarithmic phase. The expression of an outer membrane haemin-binding protein (Omp31) was up-regulated in *Brucella melitensis* during late-logarithmic phase compared to stationary phase (Rossetti *et al.*, 2009). Similarly, Davies *et al.* (1992) demonstrated changes of some OMPs (18 kDa, 24 kDa, 40.5 kDa and 94 kDa proteins) when the growth of *Mannheimia haemolytica* progressed from logarithmic phase to stationary phase. The expression of an autotransporter outer membrane adhesin AIDA-I was at a maximum level at the beginning of stationary phase and was induced by nutrient limitation in pathogenic *E. coli* (Berthiaume *et al.*, 2010). These authors suggested that an increased adherence in response to nutrient starvation could enhance survival chances. To support this, Walker *et al.* (2005) showed that *E. coli* cells at stationary phase were more adhesive than at the log phase. Another study found that the highly abundant OmpA was down-regulated when *E. coli* enter stationary phase by an increased expression of a small non-coding RNA SraD which decreased stability of the *ompA* transcript (Rasmussen *et al.*, 2005).

4.1.3.2 Different aeration rates

Bacteria can be broadly classified based on oxygen requirements into aerobes, anaerobes and facultative anaerobes (Cowan, 2012). Aerobic bacteria are able to utilize oxygen in their metabolism and process toxic oxygen products (singlet oxygen and hydrogen peroxide). If aerobic bacteria do not always require oxygen for their metabolisms and are able to grow in the absence of oxygen, these bacteria are called facultative anaerobes. Bacteria that cannot tolerate oxygen, because they lack the metabolic enzymes required for oxygen utilization and detoxification, are called anaerobes. The effect of different aeration rates on the expression of OMPs was studied by Davies *et al.* (1992) who cultured *M.*

haemolytica under anaerobic conditions, under various aerobic conditions, and with aeration in the presence of 5% CO_2. These authors demonstrated decreased expression of iron-regulated OMPs when aeration rates increased and reduced expression of a 40.5 kDa protein under anaerobic conditions.

4.1.3.3 Iron limitation

Iron is an important nutrient for bacterial cellular processes, e.g. electron transport, synthesis of amino acids, nucleosides and DNA, peroxide reduction, oxygen transport and photosynthesis. However, the availability of iron to bacteria in most environments is limited (Wandersman & Delepelaire, 2004). Therefore, bacteria have various iron acquisition mechanisms to sequestrate iron from different sources. Sources of iron can be classified into direct and indirect sources (Wandersman & Delepelaire, 2004).

First, bacteria can acquire iron directly from different sources (Ekins *et al.*, 2004; Wandersman & Delepelaire, 2004; Krewulak & Vogel, 2008). The soluble Fe(II) form can diffuse through outer membrane porin proteins under anaerobic and reducing conditions. Iron-glycoprotein complexes are found in serum (transferrin) and in lymph and mucosal secretions (lactoferrin). Another direct iron source is haem, an iron-protoporphyrin complex, but haem is highly toxic and the presence of free haem is rare. Therefore, bacteria acquire haem molecules from various haem-protein complexes e.g. haemoglobin (haem-globin complex found in blood), haptoglobin-haemoglobin (found in serum), haemopexin (haem-haemopexin complex found in plasma) and albumin (found in serum). Certain outer membrane iron receptors are responsible for direct binding to specific iron compounds. These include a bipartite transferring-binding protein TbpAB, a bipartite lactoferrin-binding protein LbpAB, a haem and haemoglobin-binding protein HemR, a haemoglobin-binding protein HmbR and a haptoglobin-haemoglobin-binding protein HgbA.

Second, bacteria can acquire iron indirectly by secreting compounds into the extracellular environment that sequestrate iron; these compounds include siderophores and haemophores (Faraldo-Gómez & Sansom, 2003; Wandersman & Delepelaire, 2004; Cescau *et al.*, 2007; Miethke & Marahiel, 2007; Krewulak & Vogel, 2008; Sandy & Butler, 2009). Siderophores bind to iron with high affinity. Once bound to iron, specific outer membrane siderophore receptors internalize the iron-siderophore complex. Examples of siderophore receptors include the enterobactin receptor FepA, the ferrichrome receptor FhuA and the ferric dicitrate receptor FecA. Similarly, haemophores are secreted to scavenge haem or haem complexes. In *H. influenzae*, the haemophore HxuA binds to the haem-haemopexin complex and delivers haem to an outer membrane receptor HxuC. Another haemophore, HasA, occurs in several bacteria and binds to haem, or releases haem from haemoglobin, and presents it to an outer membrane receptor HasR.

Once bound to the outer membrane receptors, small compounds such as siderophores and haem can be transported directly across the outer membrane. Haem or iron has to be cleaved from large complexes such as transferrin, haemoglobin and haemophore. before being transported into the periplasmic space (Wandersman & Delepelaire, 2004; Krewulak & Vogel, 2008). The energy driving this process is provided by the TonB-ExbB-ExbD system. A periplasmic-binding protein will transfer the iron compounds from the outer membrane receptors to the inner membrane ATP-binding cassette (ABC) transporter proteins which then delivered them into cytoplasm (Wandersman & Delepelaire, 2004; Krewulak & Vogel, 2008).

4.1.3.3.1 Effect of iron on the growth of bacteria

As iron availability in most environmental conditions is low, bacteria have to efficiently acquire sufficient iron to maintain their iron homeostasis (Andrews *et al.*, 2003). This can be achieved by several strategies including an ability to scavenge different forms of iron at high affinity by multiple iron receptors, storage of iron when the external supplies are

plentiful, regulation of the expression of iron-containing proteins during iron-restricted conditions and the use of redox stress response systems (Andrews *et al.*, 2003). To respond to the restricted iron availability, bacterial growth can be slowed down. Growth of *Salmonella choleraesuis* (Ho *et al.*, 2004) and *Francisella tularensis* (Lenco *et al.*, 2007) is retarded under iron-limited conditions. The genes involved in iron transport were up-regulated, whereas the genes involved in iron consumption were down-regulated, under iron-deplete conditions in *A. pleuropneumoniae* (Deslandes *et al.*, 2007; Klitgaard *et al.*, 2010). Virulence genes are also up-regulated during iron starvation (Andrews *et al.*, 2003; Ho *et al.*, 2004; Xiong *et al.*, 2010).

The effects of iron limitation on the expression of OMPs have been studied in many bacteria including *Haemophilus* species (Niven *et al.*, 1989; Wedderkopp *et al.*, 1993), *P. multocida* (Paustian *et al.*, 2001), *Pseudomonas* species (Heim *et al.*, 2003), *Campylobacter jejuni* (Holmes *et al.*, 2006), *S. enterica* (Chanana *et al.*, 2006), *M. haemolytica* (Davies *et al.*, 1992; Roehrig *et al.*, 2007), *Bordetella pertussis* (Vidakovics *et al.*, 2007), *E. coli* (Lin *et al.*, 2008), *Neisseria meningitidis* (van Ulsen *et al.*, 2009), *Avibacterium paragallinarum* (Abascal *et al.*, 2009), *Vibrio alginolyticus* (Xiong *et al.*, 2010), *Leptospira interrogans* (Lo *et al.*, 2010), *A. pleuropneumoniae* (Klitgaard *et al.*, 2010), *Acinetobacter baumannii* (Nwugo *et al.*, 2010), and *Yersinia pestis* (Pieper *et al.*, 2010). Siderophore receptors are normally expressed during iron starvation but are absent under iron-replete condition (Andrews *et al.*, 2003). In *S. enterica*, the expression of the same 69 kDa OMP was identified during growth under iron-limited, oxidative stress and anaerobic conditions, and was absent under normal conditions (Chanana *et al.*, 2006). Roehrig *et al.* (2007) studied gene expression of *M. haemolytica* grown under iron-limited conditions and demonstrated the up-regulation of two haemoglobin receptor genes *hmbR1* and *hmbR2* under these conditions. These authors compared the expression of these two genes and other iron responsive genes in *M. haemolytica* to those expressed in *P.*

multocida and showed that a few iron responsive or receptor genes were expressed in both bacteria suggesting different iron acquisition mechanisms used by these two bacteria having the same bovine host. Within the same host species, different strains of bacteria can express a common set of outer membrane iron receptors and also unique proteins in response to iron-limited conditions. This is supported by the work of Klitgaard *et al.* (2010) who demonstrated the expression of a common siderophore receptor CirA in all strains tested of *A. pleuropneumoniae* and three specific haptoglobin-haemoglobin receptors HpuB in moderately and highly virulent strains.

4.1.3.3.2 Iron-regulated pathways

A lack of iron can inhibit bacterial growth while an over-accumulation of this element can be toxic to bacteria because the reduced form of iron will react with oxygen, resulting in hydroxyl radicals which damage most biomolecules (Wandersman & Delepelaire, 2004). Therefore, regulation of the expression of iron-responsive and iron-containing proteins is essential to maintain iron homeostasis (Massé & Arguin, 2005). In many bacterial species the ferric-uptake regulator protein (Fur) controls the expression of iron-dependent genes. Under iron-replete condition, Fur together with iron as a prosthetic group will repress the expression of iron-regulated genes by binding to the Fur box which is located upstream of the iron-repressed genes (Andrews *et al.*, 2003). Conversely, a lack of iron will release the Fur protein, allowing the iron-repressed genes to be expressed (Andrews *et al.*, 2003).

4.1.3.3.3 Iron chelators

Iron chelators are small compounds which can strongly bind to iron (Heli *et al.*, 2011). Naturally, free iron can be chelated by host-produced chelators (e.g. transferrin, ovotransferrin and lactoferrin) which makes iron unavailable to bacteria or by bacterial-produced chelators (e.g. siderophores and ferritin). Similarly, synthetic iron chelators have been used to reduce the concentration of iron within culture media or in clinical therapy (Heli *et al.*, 2011). Examples of iron chelators include 2,2′-dipyridyl, desferrioxamine

(Desferal), and ethylenediamine-di-O-hydroxylphenyl acetic acid (EDDA). These compounds have been widely used to create *in vitro* iron-restricted conditions (Davies *et al.*, 1992; Wedderkopp *et al.*, 1993; Jacques *et al.*, 1994; Ebanks *et al.*, 2004; Cole *et al.*, 2006; Holmes *et al.*, 2006; Roehrig *et al.*, 2007; Najimi *et al.*, 2008; Klitgaard *et al.*, 2010; Eijkelkamp *et al.*, 2011; Vinckx *et al.*, 2011). Davies *et al.* (1992) compared the expression of *M. haemolytica* OMPs in the presence of natural (ovotransferrin and bovine transferrin) and synthetic (2,2′-dipyridyl, Desferal and EDDA) chelators. These authors demonstrated that 71 and 100 kDa OMPs were induced in the presence of all iron chelators, but a 77 kDa protein was induced only in the presence of synthetic chelators. This finding suggested that these iron chelators could have different iron binding mechanisms. This was supported by Ho *et al.* (2004) who observed different growth inhibition patterns of *S. choleraesuis* in the presence of a hydrophobic ferrous chelator (2,2′-dipyridyl) and a hydrophilic ferric chelator (EDDA). 2,2′-dipyridyl tended to limit maximum growth yield during stationary phase, whereas EDDA prolonged the duration of lag phase.

4.1.3.4 Different animal sera

In animals, serum is a blood component which excludes blood cells and clotting agents. Once bacteria have entered the bloodstream, serum is a source of nutrients for *in vivo* growth but also contains immune components (e.g. complements and antibodies) which are able to prevent bacterial multiplication.

4.1.3.4.1 Serum composition

Serum contains various types of soluble proteins (Miller *et al.*, 2009). Miller *et al.* (2009) described a proteomic reference map of pig serum which contains a number of proteins, e.g. albumin, globulins, MAP (major acute phase protein), transferrin, haemopexin, haptoglobin, IgM, IgG, IgA, glycoproteins, antitrypsin, antichymotrypsin, apolipoproteins, complement C3, clustein, retinol binding protein and transthyretin. This reference map

was used to detect changes between healthy and diseased pigs (Miller *et al.*, 2009). Transferrin, haemopexin and haptoglobin are potential iron sources for bacteria whereas MAP, IgM, IgG, IgA and complement C3 will inhibit or kill bacteria.

4.1.3.4.2 Effect of serum on the growth of bacteria

Bacterial responses to growth in serum are dependent on strains (susceptible or resistant to killing) and the animal source of the serum (Diallo & Frost, 2000; Muhairwa *et al.*, 2002). Serum-sensitive avian *P. multocida* strains were killed or their growth was suppressed in complement-preserved chicken serum due to complement-mediated killing activities (Diallo & Frost, 2000). On the other hand, some serum-resistant strains grew rapidly in the same chicken serum, suggesting a possible inhibition mechanism of serum components (Diallo & Frost, 2000). Another study demonstrated that a wide range of *P. multocida* strains were more resistant to turkey serum than chicken, duck and pig sera (Muhairwa *et al.*, 2002).

The effect of serum on the expression of OMPs has been studied in various bacterial species including *Proteus mirabilis* (Futoma-Kołoch *et al.*, 2006), *E. coli* (Prasadarao *et al.*, 2002; Hari-Dass *et al.*, 2005; Tyler *et al.*, 2008) and *Leptospira interrogans* (Patarakul *et al.*, 2010). Futoma-Kołoch *et al.* (2006) demonstrated induction of OMPs during the growth of resistant strains of *P. mirabilis* in human and bovine sera. OmpA was reported to bind to serum components including serum amyloid A (SAA) protein, which is a major acute phase protein, in a large number of Gram-negative bacteria and to C4b in *E. coli* (Prasadarao *et al.*, 2002; Hari-Dass *et al.*, 2005; Tyler *et al.*, 2008).

4.1.3.5 Growth on solid surfaces

Biofilm formation is the adherence and colonization of bacterial communities to a solid surface and is a cause of many bacterial infections (Sauer, 2003). The bacterial communities form tower- or mushroom-shaped microcolonies surrounded by secreted

extracellular matrixes (Sauer, 2003). Biofilm formation has been demonstrated in many bacterial species (Jacques *et al.*, 2010) including *Shewanella oneidensis* (De Vriendt *et al.*, 2005), *P. aeruginosa* (Mikkelsen *et al.*, 2007), *Porphyromonas gingivalis* (Ang *et al.*, 2008), *E. coli* (Landini, 2009), *Histophilus somni* (Sandal *et al.*, 2009), *Acinetobacter baumannii* (Shin *et al.*, 2009), *P. multocida* (Shayegh *et al.*, 2010b), *Riemerella anatipestifer* (Hu *et al.*, 2010), *N. meningitidis* (van Alen *et al.*, 2010), and *Tannerella forsythia* (Pham *et al.*, 2010). Environmental and physiological stresses stimulate biofilm formation in *E. coli* by the induction of adhesion and colonization factors (Alves *et al.*, 2010; Landini, 2009).

4.1.3.5.1 The difference between planktonic growth and biofilms

Bacterial cells growing in liquid environments are referred to as planktonic growth condition, whereas biofilm conditions refer to the formation of microcolonies on solid surfaces. Biofilm conditions promote antibiotic, host immune and stress resistance compared to the planktonic growth conditions (Coenye, 2010; Hu *et al.*, 2010). It remains unclear whether bacterial colonies grown on agar plates can be considered as biofilms due to a similar formation of microcolonies. This assumption was contradicted in the study by Mikkelsen *et al.* (2007) who compared protein expression in *P. aeruginosa* cells from planktonic growth, colonies on agar plates and biofilms . These authors showed that protein profiles of cells from colonies were similar to those of cells from planktonic growth; the profiles of cells from biofilms were similar to those of planktonic cells grown at exponential phase. However, a number of studies have demonstrated differences in gene or protein expression between planktonic and biofilm growth (Trémoulet *et al.*, 2002; De Vriendt *et al.*, 2005; Ang *et al.*, 2008; Shin *et al.*, 2009; Pham *et al.*, 2010).

The effects of biofilm formation on the expression of OMPs have been examined in several bacterial species as described below. Comparison of OMP expression in nutrient-rich and nutrient-deficient planktonic and biofilm cells of *P. multocida* showed differential

expression levels of certain OMPs between the nutrient-rich planktonic and biofilm conditions e.g. 89, 80, 83, 50, 46, 45, 36, 33, 31, 27, 17 and 14 kDa OMPs (Arun & Krishnappa, 2004). The names of these proteins were not identified by these authors. De Vriendt *et al.* (2005) compared protein expression among planktonic- and biofilm-grown *S. oneidensis* and found 59 differentially expressed proteins including a TolC-like protein AggA. Shin *et al.* (2009) performed a similar proteomic comparison of *A. baumannii* grown under planktonic and biofilm conditions and revealed changes in the expression of 23 proteins including one outer membrane iron receptor, OmpW and OprE3. Outer membrane iron receptors were also up-regulated in biofilm-grown *T. forsythia* (Pham *et al.*, 2010) and *P. gingivalis* (Ang *et al.*, 2008) compared to planktonic-grown cells. Ang *et al.* (2008) also demonstrated differential expression of other integral OMPs and outer membrane lipoprotein Omp28 when these two growth conditions were compared in *P. gingivalis*. Certain OMPs have roles in biofilm formation. OmpA was found to be involved in biofilm formation in *E. coli* (Orme *et al.*, 2006; Ma & Wood, 2009b) and *A. baumannii* (Gaddy *et al.*, 2009). Another important example is the *tad* locus which encodes three OMPs (RcpA, RcpB and TadD) required for biofilm formation in *A. actinomycetemcomitans* (Perez *et al.*, 2006).

4.1.3.5.2 Mechanism and test of biofilm formation

Biofilm formation is a multi-step process (Jacques *et al.*, 2010). First, bacteria have to attach to the surface by using surface components. Once attachment occurs, bacteria aggregate into microcolonies. In the next step, biofilms begin to form as the attached bacteria grow and divide. During this stage, extracellular matrices are secreted to cover and bind the bacterial microcolonies, forming flat, mushroom-like or tower-like shapes. These bacterial communities can sense and adapt to environmental change. Finally, with the induction by several stimuli, the bacteria will detach and disperse from their former colonies and initiate attachment to a new surface.

Various methods have been used to test bacterial ability of biofilm formation. *In vitro* and *in vivo* systems to mimic the formation of biofilms were extensively reviewed by Coenye and Nelis (2010). The *in vitro* systems mostly include microtiter plate-based systems, flow displacement systems in which growth media can be circulated and cell-culture-based systems (Coenye & Nelis, 2010). The *in vivo* systems use insertions of a variety of biomaterials into animal models. Biofilm-forming ability can also be tested by growing bacteria on Congo red agar plates which detect the production of extracellular matrices (Abdallah *et al.*, 2009; Eroshenko *et al.*, 2010). It was shown in a Gram-positive bacterium, *Staphylococcus aureus,* that positive biofilm-producing strains had black colonies, whereas negative strains produced red colonies (Mariana *et al.*, 2009). Jain and Agarwal (2009) demonstrated that Congo red agar was a reliable method to test biofilm formation in *S. aureus*. A study in *A. actinomycetemcomitans* demonstrated that rough-phenotype strains produced amyloid-like fibers which bound to Congo red dye within the agar forming dark red colonies, whereas negative strains produced white or opaque red colonies (Kimizuka *et al.*, 2009). Knobloch *et al.* (2002) evaluated different detection methods of biofilm formation in *S. aureus* and demonstrated that tube test correlated well with microtiter-plate methods compared to the Congo red agar method. The tube test stains attached bacterial colonies at the bottom of test tubes by crystal violet after removal of all media and free-floating cells (Knobloch *et al.*, 2002).

4.1.4 Objectives

This chapter aimed to compare the expression of OMPs in eight representative isolates of *P. multocida* associated with different diseased animal hosts under different growth conditions using a combination of proteomic techniques. The aim was understand the responses of individual isolates to different growth conditions and to compare the responses of all isolates to the same growth conditions by observing differential expressions of OMPs. These growth conditions included different stages of the growth,

different rates of aeration, iron limitation, growth in different sera alone and in combination with various culture media and growth on solid surface as biofilms.

4.2 Materials and methods

4.2.1 Bacterial isolates and growth conditions

The isolates were stored at -80ºC in 50% (v/v) glycerol in BHIB. From -80ºC stock cultures, bacteria were streaked onto BHI blood agar and incubated overnight at 37ºC. For preparation of outer membranes, the isolates were cultured overnight in 10 ml volumes of BHIB at 37ºC on an orbital shaker at 120 rpm. An overview of all methodologies for growing isolates of *P. multocida* under different growth conditions is described in **Figure 4-1**.

4.2.1.1 Different growth stages

Bovine isolate PM632 and porcine isolate PM684 of *P. multocida* were selected for these studies. Eight hundred microlitre volumes of overnight cultures were inoculated into pre-warmed 400 ml volumes of BHIB in 2-litre non-dimpled Erlenmeyer flasks and incubated at 37ºC on an orbital shaker at 120 rpm. These isolates were harvested at four different growth stages, namely mid-log phase (5-6 h), late-log phase (6-7 h), early stationary phase (8-9 h) and late stationary phase (24 h) (**Figure 4-1**).

4.2.1.2 Different aeration rates

Bovine isolate PM632 and porcine isolate PM684 of *P. multocida* were selected for these studies. The isolates were grown under three different aeration rates, namely no, normal and high rates of aeration (**Figure 4-1**). Eight hundred microlitre volumes of overnight were inoculated into 2-litre non-dimpled Erlenmeyer flasks overlayed with mineral oil for no aeration; into pre-warmed 400 ml volumes of BHIB in 2-litre non-dimpled Erlenmeyer flasks for normal aeration; and into 2-litre dimpled flasks for high aeration. The cultures were incubated until the $OD_{600\ nm}$ reached 0.8-0.9 (late-log phase) at 37ºC on an orbital

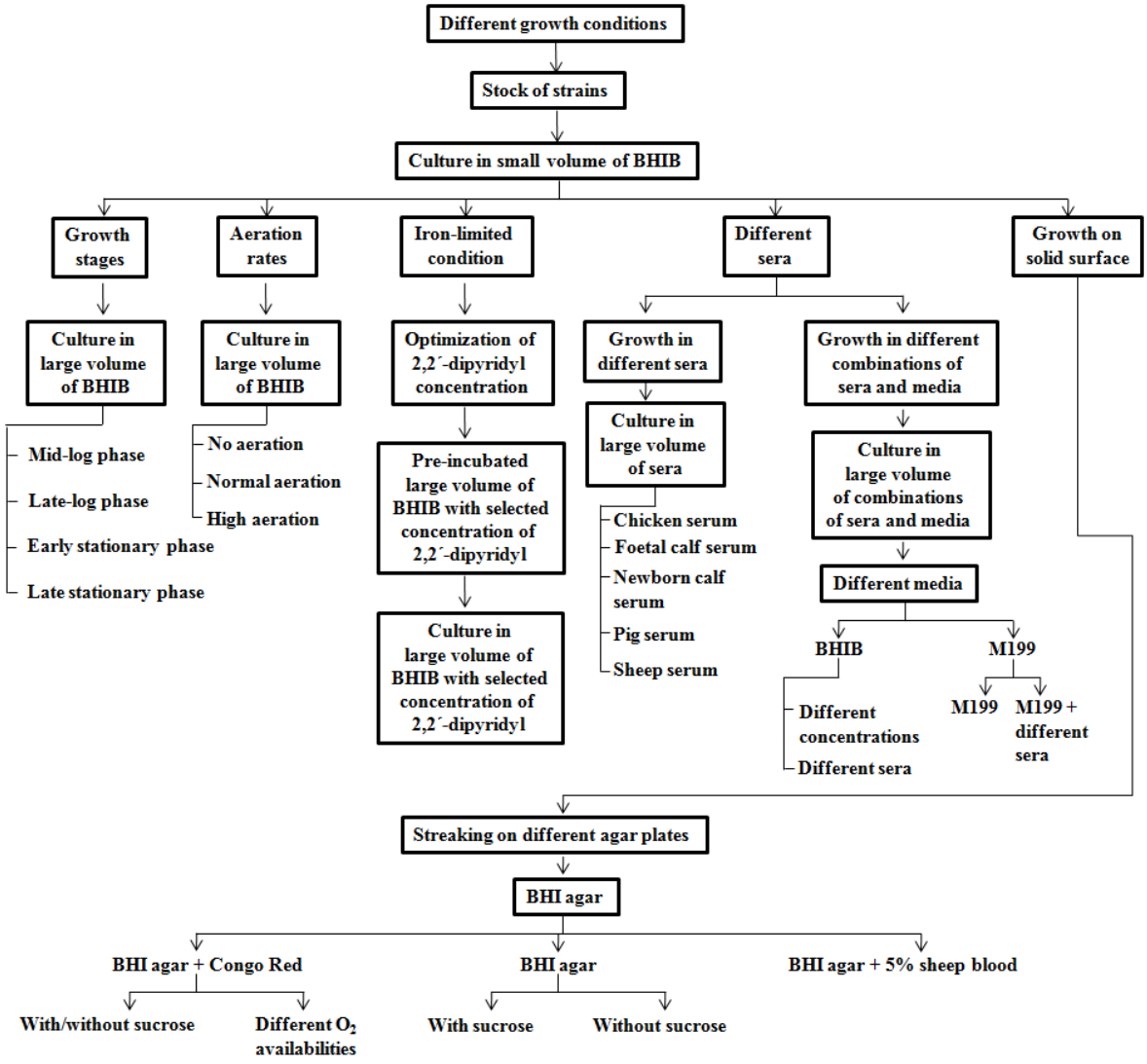


Figure 4-1. Overview of all methodologies for growing isolates of *P. multocida* under different growth conditions. Five growth conditions were examined in the present study. These included different growth stages, different aeration rates, iron limitation, different sera and growth on solid surfaces.

shaker at 120 rpm for the normal and high aeration conditions, and at 37ºC without shaking for the no aeration condition.

4.2.1.3 Iron limitation

Eight representative isolates of *P. multocida* were selected for these studies. A concentration of the iron chelator 2,2′-dipyridyl was first optimized for each isolate (**Figure 4-1**). This concentration reduced bacterial growth compared to growing bacteria in iron-rich medium but did not kill them. Fifty microlitre volumes of each isolate was inoculated into pre-warmed 25 ml volumes of BHIB in 100-millilitre non-dimpled flasks supplemented with different concentrations (0-250 μM) of 2,2′-dipyridyl . The cultures were incubated for 12 h at 37ºC on an orbital shaker at 120 rpm. Bacterial growth was monitored at every hour by reading the $OD_{600\ nm}$. Next, 0.8 ml volumes of overnight culture of each of the eight isolates were inoculated into pre-warmed 400 ml volumes of BHIB supplemented with the selected concentrations of 2,2′-dipyridyl (for each individual isolate) in 2-litre non-dimpled Erlenmeyer flasks and incubated at 37ºC on an orbital shaker at 120 rpm. The cultures were incubated until the $OD_{600\ nm}$ reached 0.8-0.9 (late-log phase).

4.2.1.4 Different animal sera

Eight representative isolates of *P. multocida* were selected for these studies. Comparative growth of these isolates in five animal sera including chicken (SLI), foetal calf (Invitrogen), newborn calf (Invitrogen), pig (SLI) and sheep (SLI) sera, was examined (**Figure 4-1**). Sera were thawed at 37ºC and heat-inactivated for 30 min at 56ºC. Thirty microlitre volumes of overnight cultures of each isolate were inoculated into heat-inactivated pre-warmed 15 ml volumes of sera in 50-millilitre non-dimpled Erlenmeyer flasks and incubated for 12 h at 37ºC on an orbital shaker at 120 rpm. Bacterial growth was monitored every two hours by reading the $OD_{600\ nm}$. For large volume cultures, 0.4 ml volumes of overnight cultures of each of these eight isolates were inoculated into heat-

inactivated, pre-warmed 200 ml volumes of sera and incubated at 37ºC on an orbital shaker at 120 rpm until the growth reached late-log phase. In this experiment, the avian isolates PM144 and PM246 were grown in chicken serum, the bovine isolates PM564 and PM632 in foetal and newborn calf sera, the porcine isolates PM684 and PM734 in pig serum and the ovine isolates PM966 and PM982 in sheep serum.

4.2.1.5 Different combinations of culture media and animal sera

The effect of different combinations of culture media and sera was first examined in the bovine isolate PM632 and the porcine isolate PM684. Four hundred microlitre volumes of overnight cultures of isolate PM632 were inoculated into pre-warmed 200 ml volumes of heat-inactivated foetal calf serum, BHIB supplemented with 10% (v/v) foetal calf serum, BHIB supplemented with 50% (v/v) foetal calf serum, M199 and M199 supplemented with 10% (v/v) foetal calf serum. Similarly, the same volumes of isolate PM684 were inoculated into pre-warmed 200 ml volumes of heat-inactivated pig serum, BHIB supplemented with 10% (v/v) pig serum, BHIB supplemented with 50% (v/v) pig serum, M199 and M199 supplemented with 10% (v/v) pig serum. These cultures were incubated at 37ºC on an orbital shaker at 120 rpm until the growth reached late-log phase.

The effect of growing eight representative isolates in a selected culture medium (M199) and in the presence or absence of sera specific to the host of origin was examined. Six hundred microlitre volumes of overnight cultures of each of the eight isolates were inoculated into pre-warmed 300 ml volumes of M199 and M199 supplemented with 10% (v/v) serum specific to the host of origin. These cultures were incubated at 37ºC on an orbital shaker at 120 rpm until the growth reached late-log phase.

4.2.1.6 Growth on solid surface or biofilm condition

The eight isolates of *P. multocida* were grown on different types of agar media in an attempt to mimic growth on a solid surface as a biofilm. The different agar media used

included BHI agar supplemented with 0.08% Congo red in the presence or absence of 5% sucrose, BHI agar alone in the presence or absence of 5% sucrose, and BHI agar supplemented with 5% defibrinated sheep blood (**Figure 4-1**). Five single colonies of overnight plate-grown of each of these eight isolates grown on BHI blood agar were resuspended in 1.5 ml volumes of PBS. One hundred microlitre volumes of the bacterial suspensions were spreaded on the surface of the agar plates and incubated for 24 h at 37ºC. Colony morphology and colour were observed and compared.

4.2.2 Preparation of OMPs

Outer membrane proteins were prepared by Sarkosyl extraction as previously described in **Chapter 3**.

4.2.3 Protein separation

The extracted outer membrane proteins were separated by 1-D SDS-PAGE as previously described in **Chapter 3**.

4.2.4 Proteomic analyses

Outer membrane protein samples were analyzed by gel-based and gel-free proteomic methods as previously described in **Chapter 3**.

4.3 Results

4.3.1 OMP profiles after different growth stages

The bovine isolate PM632 and porcine isolate PM684 were selected to examine the OMP profiles during different growth stages: mid-log phase, late-log phase, early stationary phase and late stationary phase. The OMP profiles of these two isolates were very similar at each of the four growth stages. Slight quantitative changes could be observed in a few bands such as HgbA, TolC, FadL and OmpW. These are shown in **Figure 4-2**.

4.3.2 OMP profiles after growth under different aeration rates

The bovine isolate PM632 and porcine isolate PM684 were selected to examine changes in OMP profiles under different aeration rates: normal aeration, high aeration and no aeration. Bands of differentially expressed proteins were cut out and identified by the gel-based proteomic method. The identified OMPs were identical to those identified in **Figure 3-10**. High molecular mass bands including HgbA, Oma87 and TbpA of isolate PM632 (Box A in **Figure 4-3**) were expressed at low levels after growth with high aeration compared with growth at normal and no aeration. In Box B of the isolate PM684 (**Figure 4-3**), the upper band of HgbA was expressed at lower level after growth with no aeration compared to normal and high aeration. In Box C of isolate PM632 (**Figure 4-3**), the bands of TolC and FadL (upper bands) and HexD and GlpQ (lower bands) were expressed at a lower level after growth with no aeration compared to normal and high aeration. Lower levels of expression were also observed for HexD and GlpQ in isolate PM684 (Box D in **Figure 4-3**) and for PlpA/MetQ in isolate PM632 (Box E in **Figure 4-3**) after growth with no aeration. On the other hand, OmpA and MipA/OmpV of isolate PM632 (BoxF in **Figure 4-3**) and OmpW of isolate PM684 (Box H in **Figure 4-3**) were expressed at higher levels after growth with no aeration compared to the other conditions. OmpW of isolate PM684 (Box G in **Figure 4-3**) was expressed at higher level after growth with normal aeration compared to growth with high and no aeration.

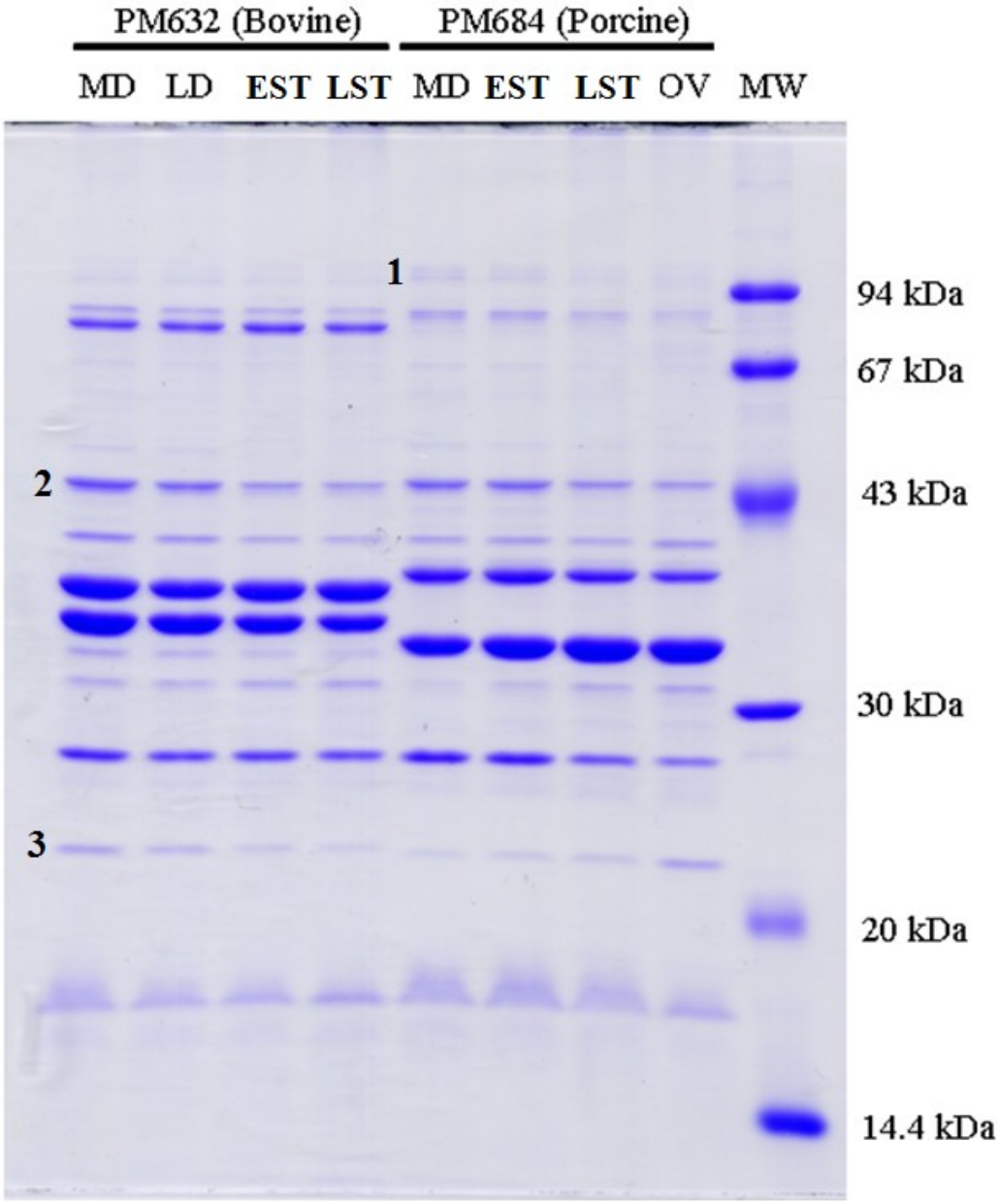


Figure 4-2. Comparison of the OMP profiles of isolates PM632 and PM684 of *P. multocida* harvested at different growth stages: mid-log phase (MD), late-log phase (LD), early stationary phase (EST) and late stationary phase (LST). Numbers indicates protein bands which were quantitatively changed in expression: 1, HgbA; 2, TolC and FadL; and 3, OmpW.

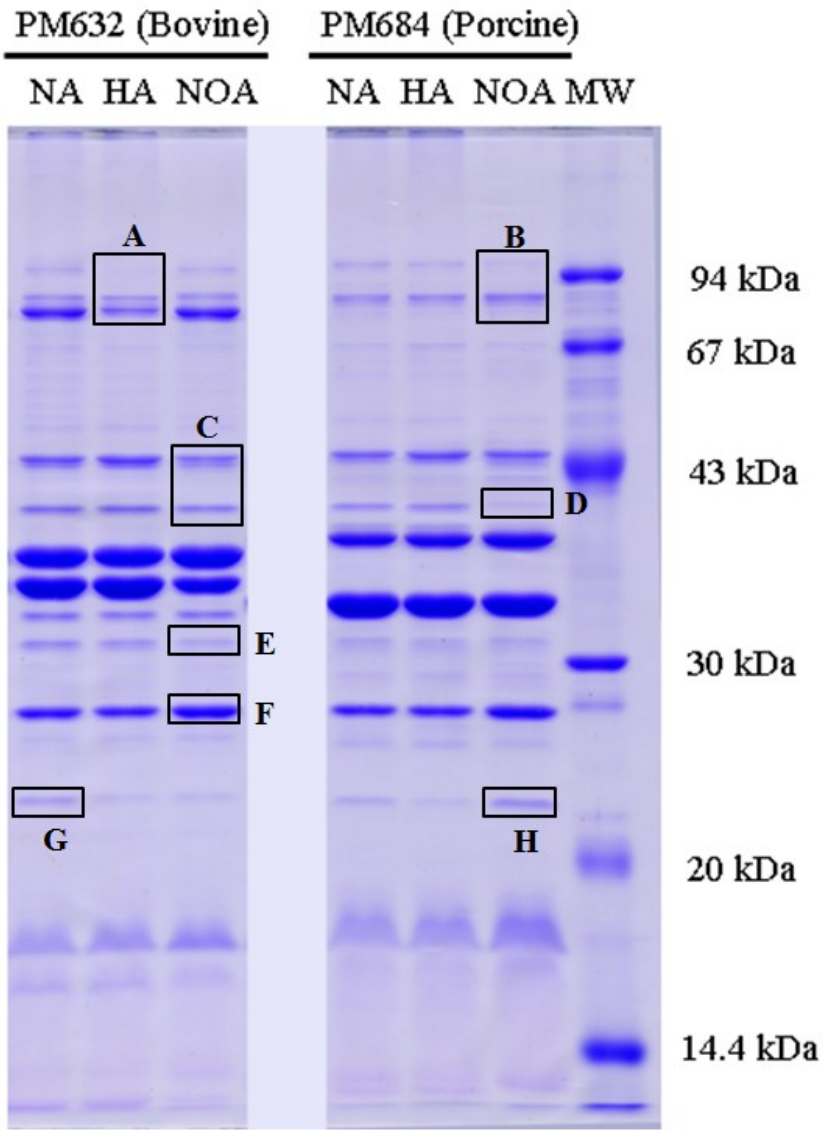


Figure 4-3. Comparison of the OMP profiles of isolates PM632 and PM684 of *P. multocida* grown under different rates of aeration: normal aeration (NA), high aeration (HA) and no aeration (NOA). Labelled boxes indicate protein bands with changed levels of expression (described in the text).

4.3.3 OMP profiles after growth under iron-limited conditions

The eight isolates of *P. multocida* were grown under iron-limited contidions. The synthetic iron chelator 2,2′-dipyridyl was used to remove iron from the growth media. These isolates were grown under a range of dipyridyl concentrations shown in **Figures 4-4 and 4-5**. Optimal concentrations of dipyridyl which reduced the growth but did not kill the bacteria were selected for individual isolates as follows: 135 μM for isolates PM144, PM246 and PM564; 100 μM for isolate PM632; 90 μM for isolate PM684; 45 μM for isolate PM734; 150 μM for isolate PM966; and 130 μM for isolate PM982. These experiments showed that increased concentrations of dipyridyl reduced the growth of *P. multocida* and the eight isolates of *P. multocida* showed different tolerance levels to iron-limited conditions.

The eight isolates of *P. multocida* were grown in media supplemented with the optimum concentration of dipyridyl. Comparisons of growth under iron-replete and iron-limited conditions are shown in **Figure 4-6**. Because the bacterial growth was decreased, the growing times to reach the same cell density had to be increased. Five isolates (PM144, PM246, PM632, PM684 and PM734) grown under iron-limited conditions reached log-phase approximately one hour later than those grown under iron-replete condition. Three isolates (PM564, PM966 and PM982) grown under iron-limited condition reached log-phase approximately five hours later than those grown under iron-replete condition. The OMP profiles prepared from bacterial cells grown under iron-limited condition are illustrated in **Figure 4-7** in comparison with those prepared from cells grown under iron-replete condition. Clearly, the major changes in the OMP profiles were found in the region between 70 to 100 kDa (**Figure 4-7**). Certain proteins were clearly indicated in this region. For example, TbpA was highly expressed in both bovine isolates and ovine isolate PM982 (No. 5 in **Figure 4-7**); HgbA was identified in all isolates and also showed

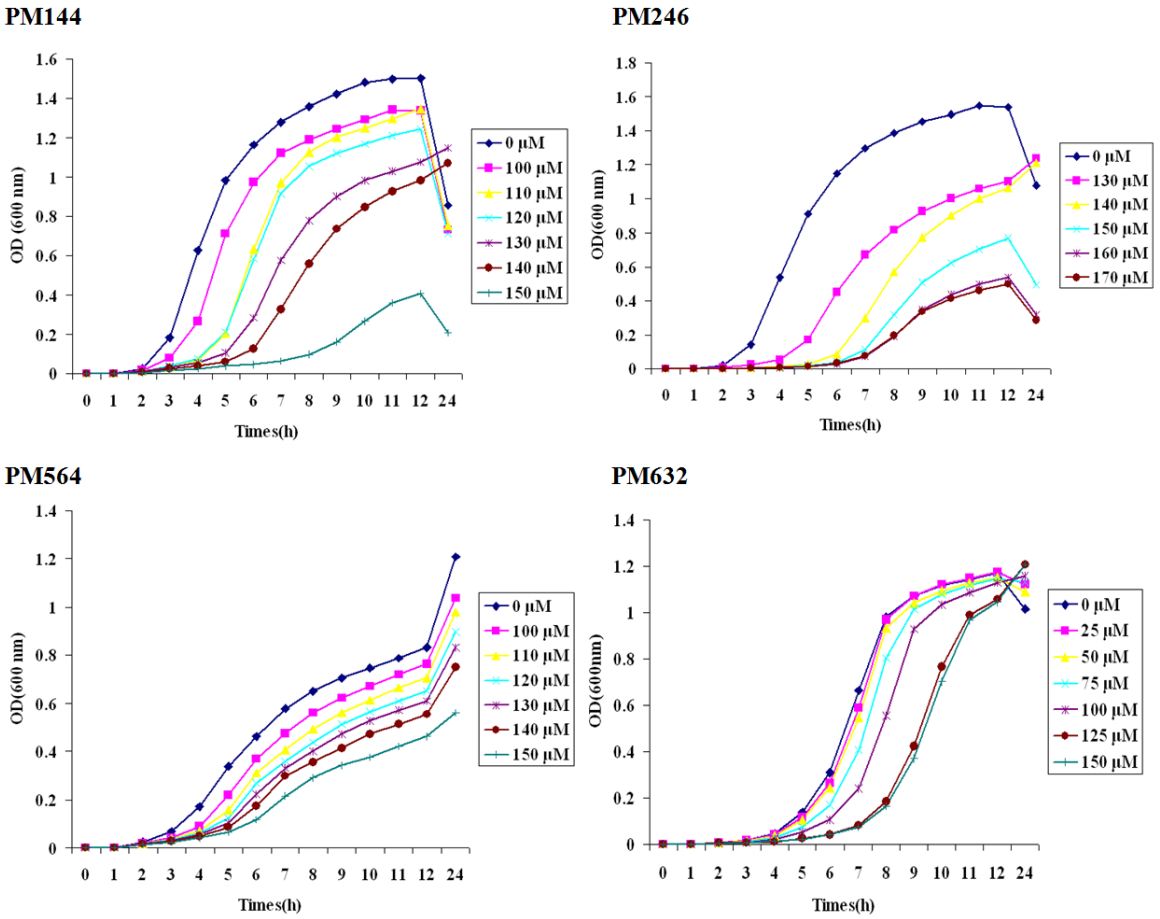


Figure 4-4. Growth of *P. multocida* isolates PM144, PM246, PM564 and PM632 under a range of dipyridyl concentrations. These isolates were grown in BHIB pre-incubated with different concentrations of 2,2´-dipyridyl (shown in the right boxes) at 37ºC. The growth was monitored by reading $OD_{600\ nm}$. The x-axis shows growth times (h) and the y axis shows $OD_{600\ nm}$. These graphs represent one replicate of results.

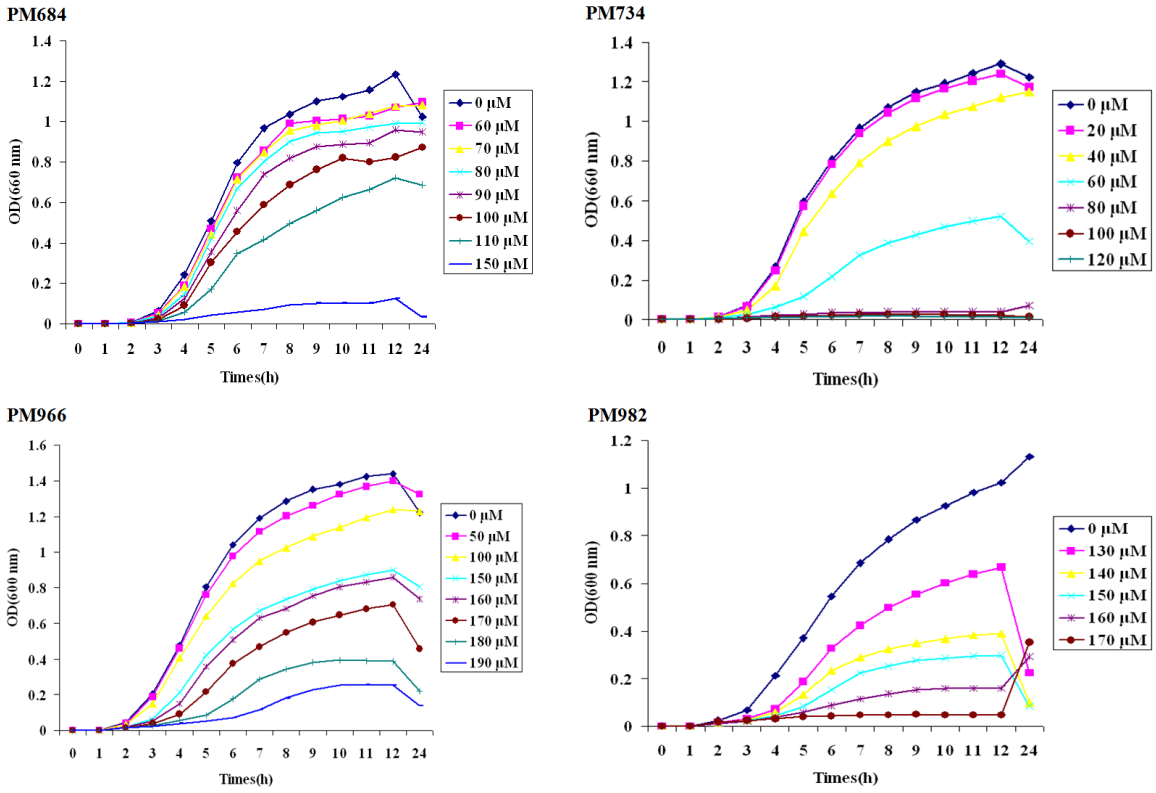


Figure 4-5. Growth of *P. multocida* isolates PM684, PM734, PM966 and PM982 under a range of dipyridyl concentrations. These isolates were grown in BHIB pre-incubated with different concentrations of 2,2′-dipyridyl (shown in the right boxes) at 37ºC. The growth was monitored by reading $OD_{600\ nm}$. The x-axis shows growth times (h) and the y axis shows $OD_{600\ nm}$. These graphs represent one replicate of results.

molecular mass variation (No.1 in **Figure 4-7**). The gel-based proteomic analysis showed that visible bands in this region contained more than one protein. These included a number of iron receptors, Oma87 and OstA.Analysis of the OMPs expressed under iron-limited conditions by the gel-based and gel-free approaches showed 46 OMPs (**Table 4-1**). Eight core OMPs were identified in all isolates including HgbA, Oma87, FadL, OmpA, OmpH_1, Pal/Omp P6, OmpW and Lpp/Pcp. Five OMPs were identified only under iron-limited conditions in comparison to the iron-replete conditions: four TonB-dependent iron receptors PfhR, Hup, PM0741 and PM1282 and a hypothetical protein PM1543. RcpA, Opa and TadD were identified only in the avian isolates and the ovine isolate PM966 after growth in both iron-replete and iron-limited conditions. NanH was identified only in the bovine isolate PM632 under both growth conditions. Similarly, HlpB was identified only in the avian isolate PM144 under both growth conditions. Twelve iron receptors were identified from the eight isolates (**Table 4-1**). These included HgbA, HemR, HasR, TbpA, two HgbB proteins, PfhR, Hup, and TonB-dependent receptors PM0803, PM1428, PM0741 and PM1282. The number of identified iron receptors increased when the bacteria were grown under iron-limited conditions compared to when they were grown under iron-replete conditions: from one iron receptor under iron-replete condition to seven iron receptors under iron-limited condition for isolate PM144; from one to eight iron receptors for isolate PM246; from two to three iron receptors for isolate PM564; from four to six iron receptors for isolate PM632; from four to nine iron receptors for isolate PM684; from seven to eight iron receptors for isolate PM734; from three to eight iron receptors for isolate PM966; and from three to seven iron receptors for isolate PM982 (**Table 4-1**). Of these twelve iron receptors, two OMPs (HgbA and HemR) were identified in all isolates grown under iron-limited condition, whereas only HgbA was identified in all isolates grown under iron-replete condition. Under iron-limited condition, certain OMPs were restricted in isolates from one, two or three animal hosts. Hup and TonB-dependent receptor PM1282 were identified in only the porcine isolates; TonB-dependent receptor

PM1428 was identified in only the porcine and ovine isolates; HgbB (PM0336) was identified in only the avian and ovine isolates; and PfhR was identified in only the avian, porcine and ovine isolates (**Table 4-1**).

4.3.4 OMP profiles after growth in different animal sera

The eight isolates of *P. multocida* were grown in five different animal sera including chicken serum, foetal calf serum, newborn calf serum, pig serum and sheep serum. The eight isolates grew differently in each serum. The two avian isolates (PM144 and PM246) and the ovine isolate PM966 grew higher in all five sera compared to the other isolates (**Figure 4-8**). Maximum growth occurred in chicken ($OD_{600\ nm}$ = 0.7) and foetal calf ($OD_{600\ nm}$ = 0.55) sera . These growths were higher than those achieved in newborn calf and pig sera (maximum $OD_{600\ nm}$ = 0.4) and sheep serum (maximum $OD_{600\ nm}$ = 0.2) (**Figure 4-8**). A comparison of individual isolates grown in these five sera is shown in **Figure 4-9**. Clearly, the isolates grew better in the chicken and foetal calf sera compared to newborn calf, pig and sheep sera.

To test the effect of serum on the expression of the outer membrane proteome, the eight isolates were grown in the serum associated with their respective animal hosts: the avian isolates were grown in chicken serum; the bovine isolates in foetal and newborn calf sera; the porcine isolates in pig serum; and the ovine isolates in sheep serum. The OMP profiles are shown in **Figure 4-10A**. The profiles were complex when compared with the profiles obtained after growth in BHIB. Thegreater complexity could be due to interference of serum proteins during the extraction process (**Appendix Table 4-1**). The protein profiles of the five animal sera are shown in **Figure 4-10B**. The serum protein profiles contained highly abundant proteins in the region between 43 kDa to 120 kDa.

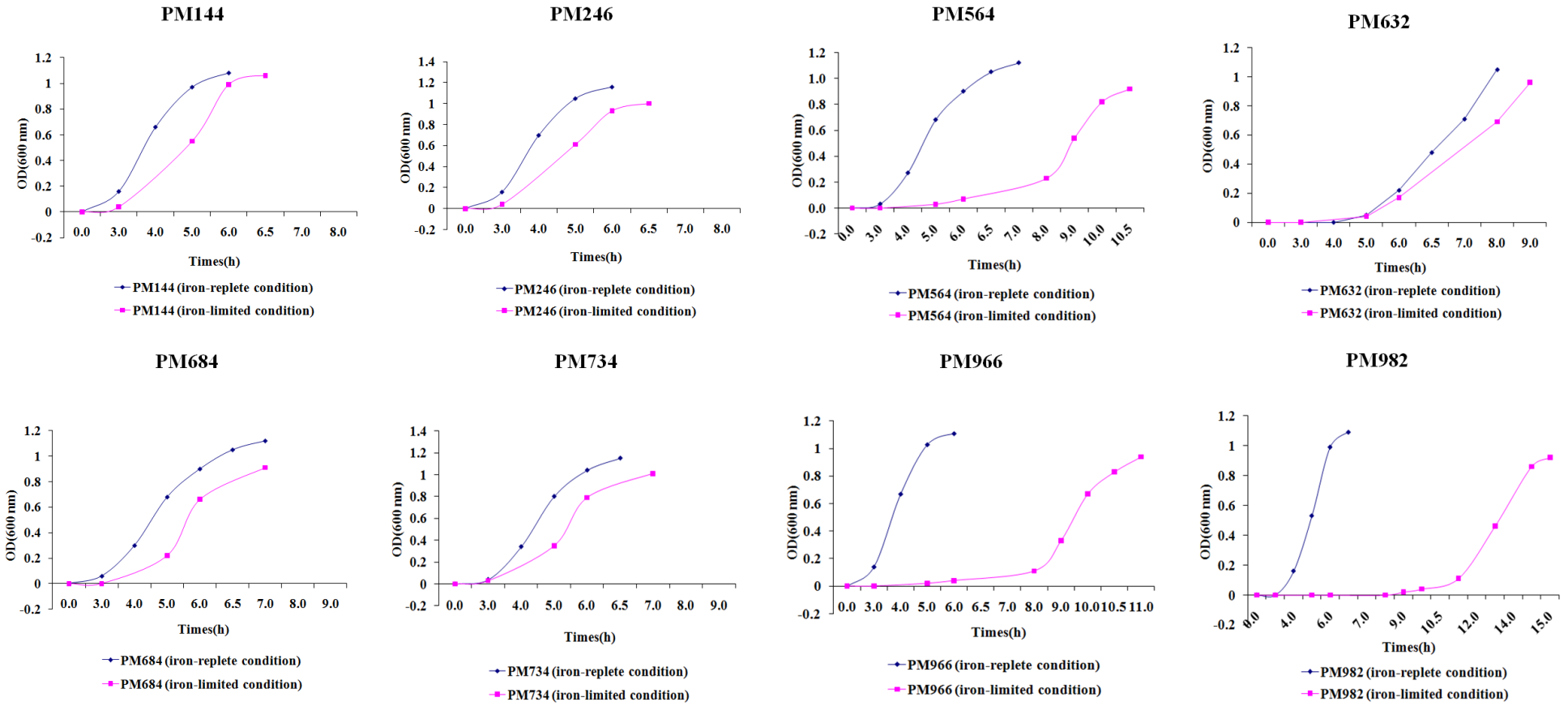


Figure 4-6. Comparison of the growth of eight isolates of *P. multocida* under iron-replete and iron-limited conditions. These graphs represent one replicate of results. The isolates were grown in BHIB for iron replete conditions and in BHIB supplemented by selected concentrations of 2,2′-dipyridyl. The x-axis shows growth times (h) and the y axis shows $OD_{600\ nm}$.

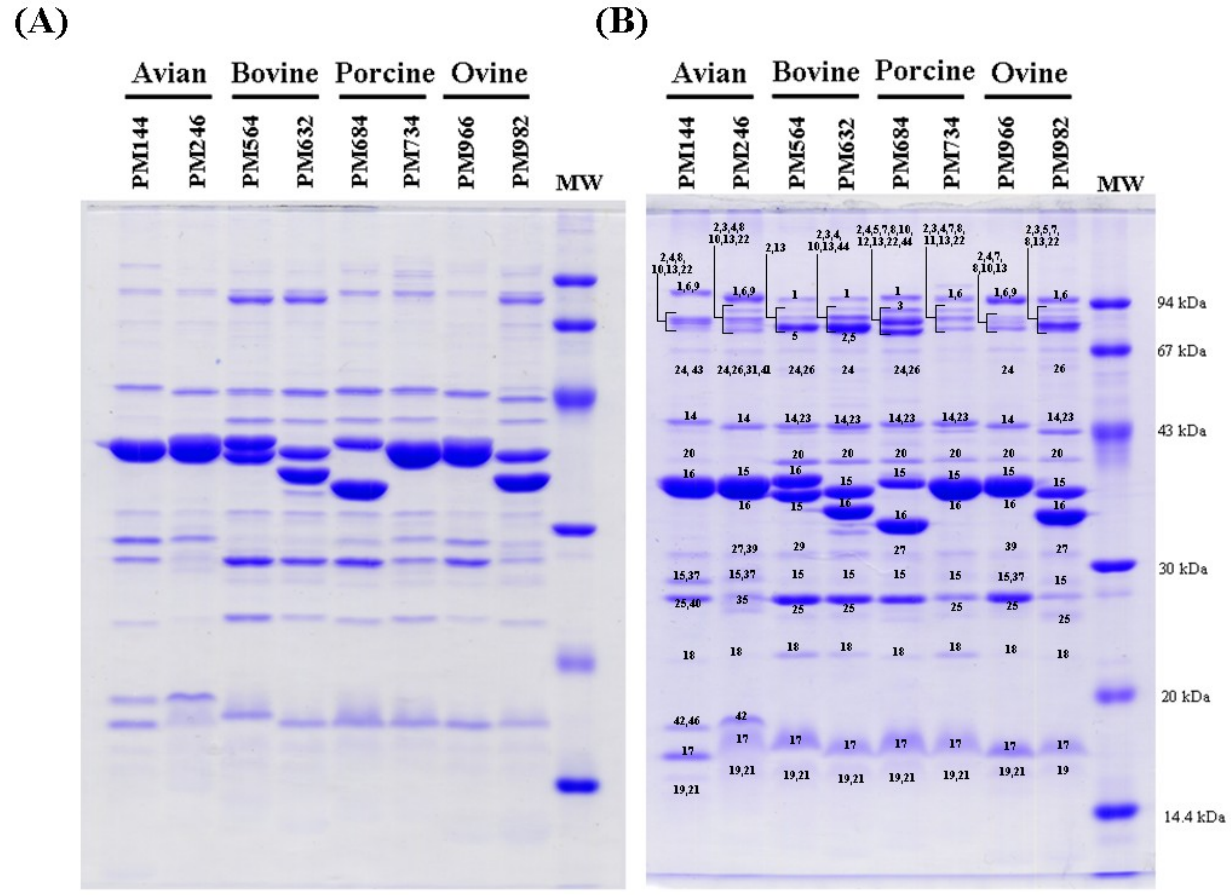


Figure 4-7. Comparison of the OMP profiles of eight isolates of *P. multocida* grown under iron-replete (A) and iron-limited (B) growth conditions. Numbers labelled on the gel correspond to the number of OMPs presented in **Table 4-1**. The gel bands were processed by in-gel trypsin digestion followed by LC-MS/MS analysis.

Table 4-1. Comparison of the OMPs identified from eight isolates of *P. multocida* grown under iron-replete and iron-limited conditions. "+" indicate the presence of that protein and "-" indicates that the protein is not observed. Protein ID and name of the iron receptors are shown in bold.

No	Protein ID	Protein name	Isolates of different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited
1	**PM0300**	**HgbA**	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2	**PM0576**	**HemR**	-	+	-	+	-	+	-	+	+	+	+	+	-	+	-	+
3	**PM1622**	**HasR**	-	-	-	+	-	-	+	+	+	+	+	+	-	-	-	+
4	**PM0803**	**TonB-dependent receptor**	-	+	-	+	-	-	+	+	-	+	+	+	-	+	-	-
5	**-**	**TbpA**	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+	+
6	**PM0337**	**HgbB**	-	+	-	+	-	-	-	-	-	-	+	+	-	+	+	+
7	**PM1428**	**TonB-dependent receptor**	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	+
8	**PM0040**	**PfhR**	-	+	-	+	-	-	-	-	-	+	-	+	-	+	-	+

Table 4-1. (Continued)

No	Protein ID	Protein name	Isolates of different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited
9	**PM0336**	**HgbB**	-	+	-	+	-	-	-	-	-	-	+	-	+	+	-	-
10	**PM0741**		-	+	-	+	-	-	-	+	-	+	-	-	-	+	-	-
11	**-**	**Hup**	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
12	**PM1282**	**TonB-dependent receptor**	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
13	PM1992	Oma87	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
14	PM1069	FadL	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
15	PM0786	OmpA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
16	PM0388	OmpH_1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
17	PM0966	Pal/Omp P6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
18	PM0331	OmpW	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
19	PM0554	Lpp/Pcp	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
20	PM0778	HexD	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+

Table 4-1. (Continued)

No	Protein ID	Protein name	Isolates of different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited
21	PM0016	Lipoprotein	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
22	PM1600	LptD/Imp/OstA	+	+	+	+	+	-	+	-	+	+	+	+	+	-	+	+
23	PM0527	TolC	+	-	+	-	+	+	+	+	+	+	+	+	+	-	+	+
24	PM1614	LppB/NlpD	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	-
25	PM0998	MipA/OmpV family protein	+	+	-	-	+	+	+	+	+	-	+	+	+	+	+	+
26	PM0076	EstA	+	-	+	+	+	+	+	-	+	+	+	-	+	-	+	+
27	PM1730	PlpA/MetQ	+	-	+	+	+	-	+	-	+	+	+	-	+	-	+	+
28	PM1050	NlpB	+	+	-	-	-	-	+	-	+	+	+	-	+	+	+	-
29	PM1064	Lipoprotein E/OmpP4	+	-	+	-	+	+	+	-	+	-	+	-	+	-	+	-
30	PM1720	ComL	+	-	+	-	+	+	-	-	-	+	+	-	+	-	+	+

Table 4-1. (Continued)

No	Protein ID	Protein name	Isolates of different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited
31	PM1809	YtfM/Omp85 family protein	+	-	+	+	+	-	-	-	+	-	+	-	+	-	+	-
32	PM1426	Phospholipase A/OmpLA	-	-	+	-	-	+	+	-	-	+	+	+	-	+	-	+
33	PM0442	Lipoprotein	+	-	+	-	+	+	+	+	-	-	+	-	+	-	-	-
34	PM1215	RlpB	-	+	+	+	-	+	-	+	+	+	-	-	-	-	-	+
35	PM0389	OmpH_2	-	+	-	+	-	+	-	-	+	+	-	+	-	+	-	-
36	PM1827	Hypothetical protein	+	-	+	+	-	-	-	-	+	-	+	-	+	+	-	-
37	PM0852	RcpA	+	+	+	+	-	-	-	-	-	-	-	-	+	+	-	-
38	PM1515	Conserved hypothetical protein	+	-	-	-	+	-	-	-	+	+	+	-	-	-	-	-
39	PM0586	Plp4	+	-	+	+	-	-	-	-	-	-	-	-	+	+	-	-
40	PM0846	TadD	+	+	+	+	-	-	-	-	-	-	-	-	+	-	-	-

Table 4-1. (Continued)

No	Protein ID	Protein name	Isolates of different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited	Iron-replete	Iron-limited
41	PM0056	LspB_1	-	-	+	+	+	-	-	-	-	-	-	-	+	-	-	-
42	PM1025	Opa	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
43	PM0058	LspB_2	-	+	+	-	-	-	-	-	-	-	-	-	+	-	-	-
44	PM0663	NanH	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
45	PM1543	Hypothetical protein	-	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-
46	PM1077	HlpB	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Total number of proteins	28	27	25	30	23	22	23	21	26	29	30	22	29	26	21	23

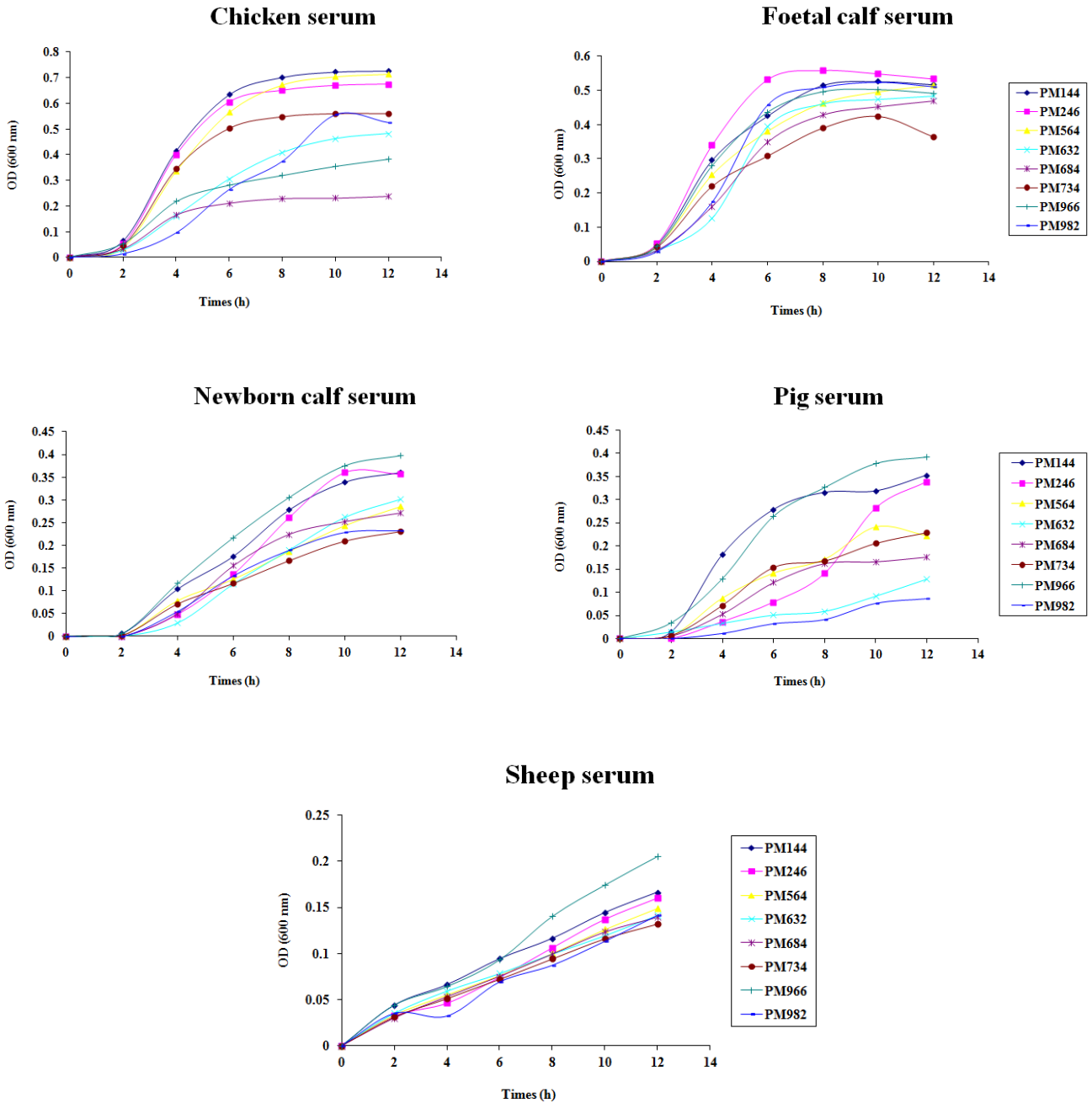


Figure 4-8. Comparison of the growth of eight isolates of *P. multocida* in five heat-inactivated animal sera: chicken, foetal calf, newborn calf, pig and sheep sera. The x-axis shows growth times (h) and the y axis shows $OD_{600\ nm}$.

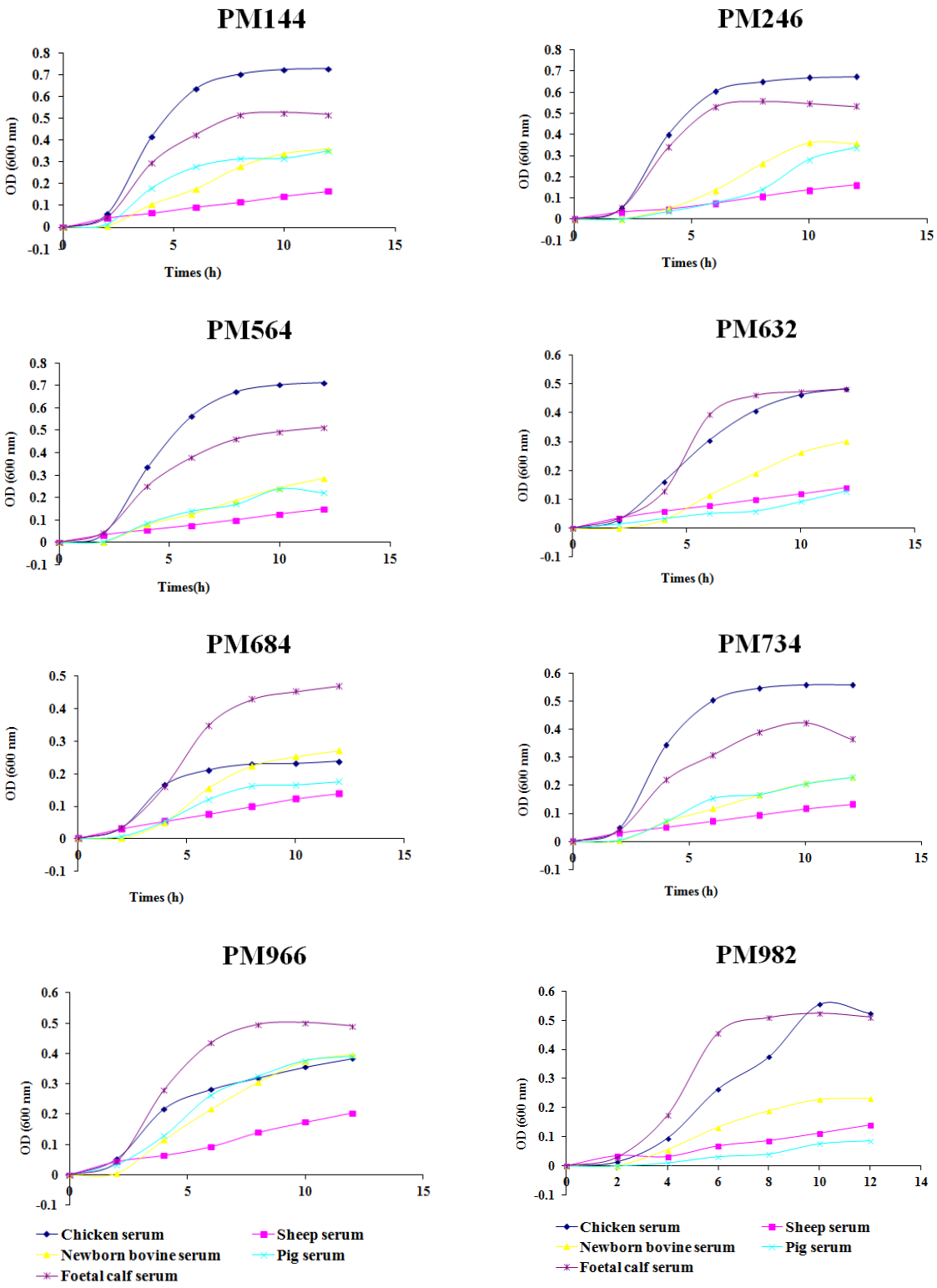


Figure 4-9. Comparison of the growth of each isolate of *P. multocida* in five animal sera: chicken, foetal calf, newborn calf, pig and sheep sera. The x-axis shows growth times (h) and the y axis shows $OD_{600\ nm}$.

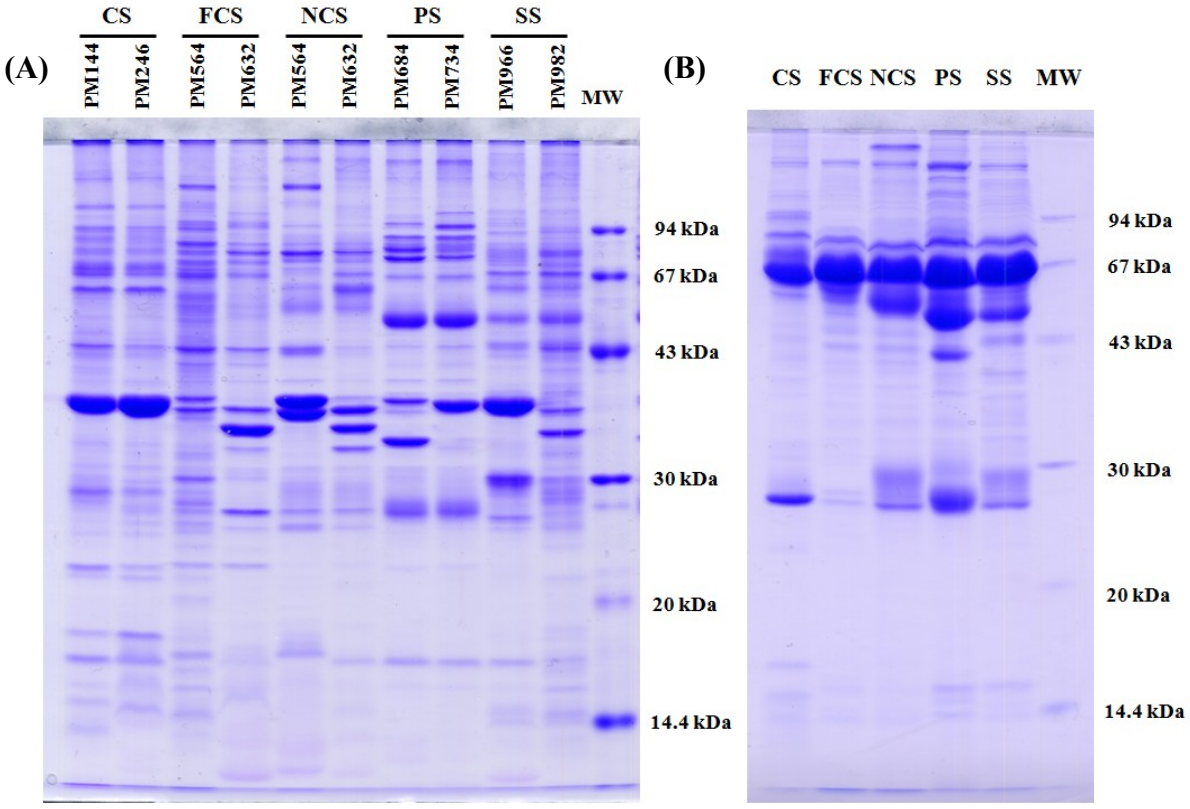


Figure 4-10. OMP profiles of eight isolates of *P. multocida* grown in serum associated with their animal hosts **(A)**. Serum protein profiles **(B)**. CS, chicken serum; FCS, foetal calf serum; NCS, newborn calf serum; PS, pig serum; and SS, sheep serum.

Growing the same two isolates of *P. multocida* (PM632 and PM684) in sera associated with their animal hosts and in BHI and M199 supplemented with different concentrations of the same sera was examined. The comparison of the OMP profiles is shown in **Figure 4-11**. For the bovine isolate PM632, the OMP profiles resulting from growth in BHIB, foetal calf serum, BHIB supplemented with 10% and 50% foetal calf serum, and M199 supplemented with 10% foetal calf serum were similar. However, the 67 kDa – 97 kDa OMPs were highly expressed when this isolate was grown in M199 alone. For the porcine isolate PM684, the OMP profile resulting from growth in BHIB was similar to that resulting from growth in BHIB supplemented with 10% pig serum (**Figure 4-11**). The 67 kDa – 97 kDa OMPs were highly expressed when this isolate was grown in pig serum, BHIB supplemented with 50% pig serum and M199 supplemented with 10% pig serum.

The effect of growing the eight isolates in M199 and M199 supplemented with 10% serum associated with their hosts of origin was examined. The OMP profiles are shown in **Figure 4-12**. The OMP profiles of the avian isolates showed high expression of the 67 kDa – 70 kDa OMPs when grown in M199 and the expression reduced when 10% chicken serum was added to the medium. The OMP profiles of the bovine isolate PM564 showed slight differences when grown in M199 compared to M199 supplemented with 10% foetal calf serum. However, the other bovine isolate PM632 produced highly expressed OMPs in the 67 kDa – 79kDa range when grown in M199, but expression of these proteins was dramatically reduced with the addition of 10% foetal calf serum. The OMP profiles of the two porcine isolates were similar when grown in M199 with and without the addition of 10% pig serum. The OMP profiles of the two ovine isolates PM966 and PM982 showed slightly higher levels of expression of proteins in the 67 kDa – 97 kDa range when grown in M199 supplemented with 10% sheep serum compared to M199 alone.

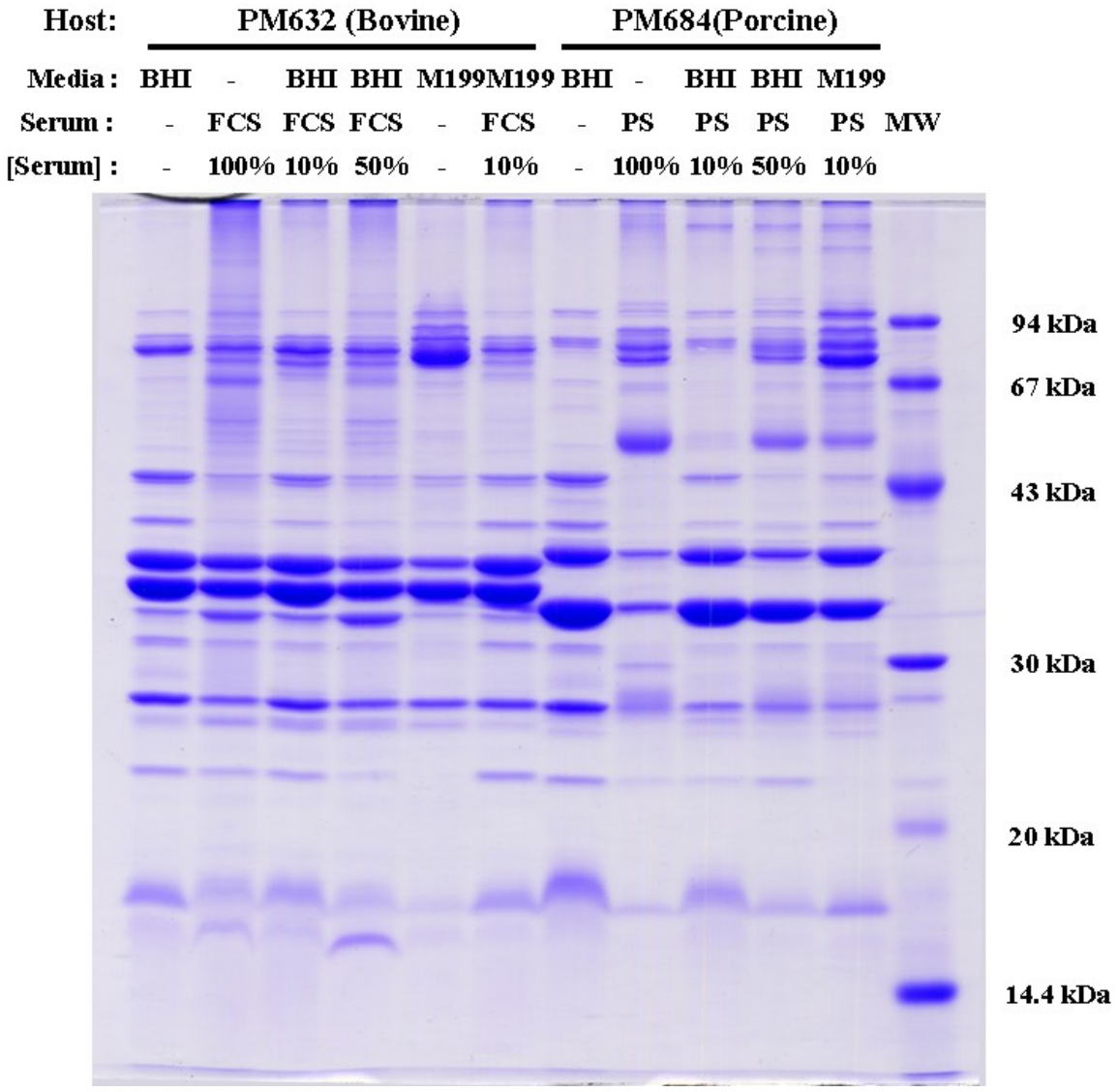


Figure 4-11. Comparison of OMP profiles of bovine isolate PM632 and porcine isolate PM684 grown in sera associated with their animal hosts and in BHIB and M199 supplemented with different concentrations of the same sera. FCS, foetal calf serum; PS, pig serum.

The OMP profiles of the eight isolates grown in BHIB, BHIB supplemented with dipyridyl (iron limitation), serum associated with their animal hosts, M199 and M199 supplemented with 10% of the same serum were compared (**Figures 4-13 to 4-16)**. For the two avian isolates PM144 and PM246 (**Figure 4-13**), the OMP profiles of cells grown under iron-limited conditions and in M199 were similar. The high expression levels of the 67 kDa-97 kDa OMPs under these two growth conditions was clear when compared with the growth of these isolates in BHIB. The OMP profiles resulting from growth in chicken serum and M199 supplemented with 10% chicken serum were more complex. Proteomic analysis of these additional bands showed that they were contaminant cytoplasmic, inner membrane, periplasmic and serum proteins (**Appendix Table 4-2**).

For the bovine isolate PM564 (**Figure 4-14**), the OMP profiles resulting from growth in BHIB, BHIB supplemented with dipyridyl, newborn bovine serum and M199 were similar. The OMP profiles resulting from growth in foetal calf serum and M199 supplemented with 10% foetal calf serum were complex and many of these additional bands were contaminant cytoplasmic, inner membrane, periplasmic and serum proteins. For the other bovine isolate PM632 (**Figure 4-14**), the increased level of expression of proteins in the 67 kDa – 97 kDa range was clearly observed this isolate was grown in BHIB supplemented with dipyridyl and in M199 compared to growth in BHIB alone. These patterns were not observed when the same isolate was grown in newborn calf serum, foetal calf serum and M199 supplemented with 10% foetal calf serum. For the two porcine isolates PM684 and PM734 (**Figure 4-15**), the OMP profiles were similar when the isolates were grown in BHIB supplemented with dipyridyl, pig serum, M199 and M199 supplemented with 10% pig serum. Expression of proteins in the 67 kDa-97 kDa range was higher in these conditions compared to growth in BHIB alone. For the ovine isolate PM966 (**Figure 4-16**), the OMP profiles resulting from from growth of this isolate in BHIB supplemented with dipyridyl, M199 and M199 supplemented with 10% sheep serum were similar. The

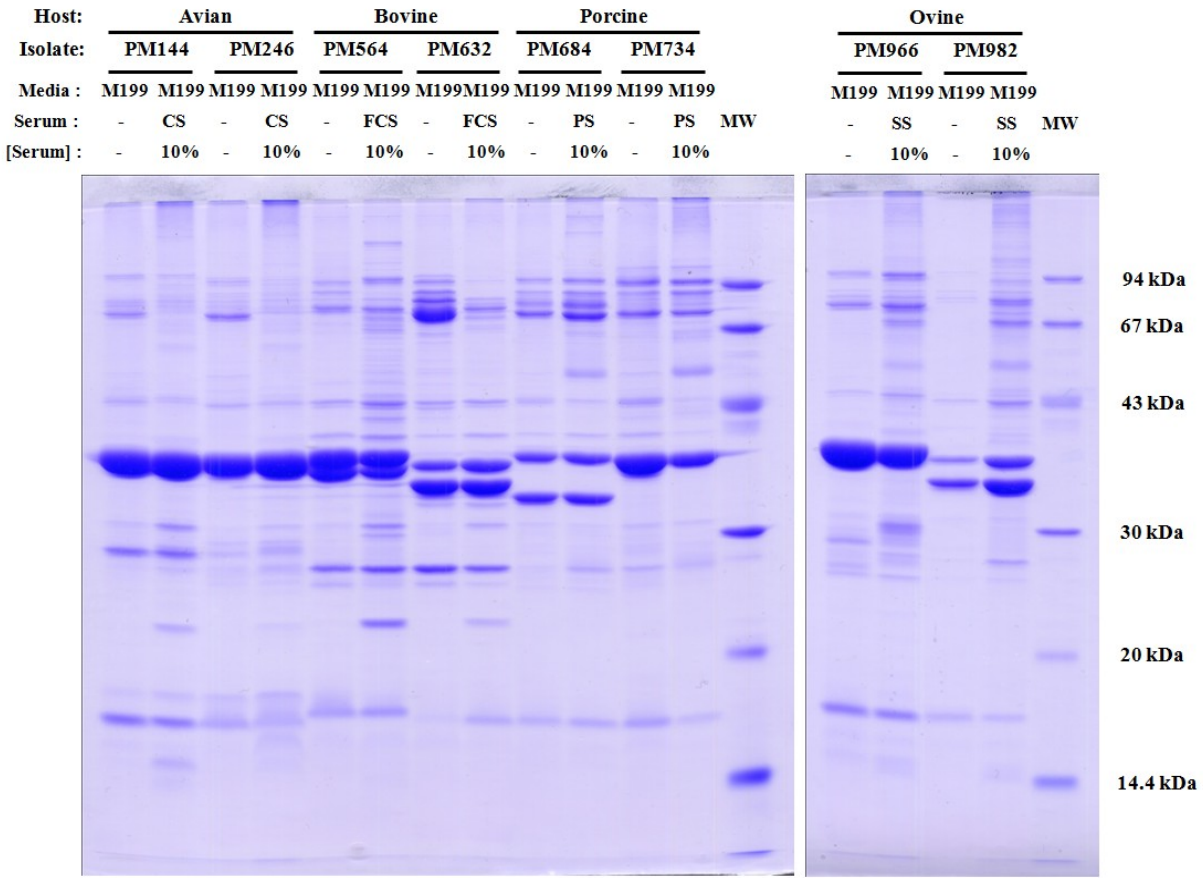


Figure 4-12. Comparison of the OMP profiles of eight isolates of *P. multocida* grown in M199 and M199 supplemented with serum associated with their hosts of origin. CS, chicken serum; FCS, foetal calf serum; PS, pig serum; and SS, sheep serum.

expression of proteins in the 67 kDa-97 kDa range was higher in these conditions compared to growth in BHIB alone. For the other ovine isolate PM982 (**Figure 4-16**), the OMP profiles resulting from growth in BHIB and BHIB supplemented with dipyridyl were similar, with slightly higher expression of proteins in the 67 kDa – 97 kDa range under the latter conditions. However, the OMP profiles resulting from growth in sheep serum and M199 supplemented with 10% sheep serum were more complex and many of these additional bands were contaminant cytoplasmic, inner membrane, periplasmic and serum proteins as shown in other isolates. Lastly, growth of isolate PM982 in M199 was very low and there was not enough outer membrane material for the comparison. Overall, there were similarities of the OMP profiles after growth under iron limited conditions and in M199 as well as in serum for certain isolates. The pig and sheep sera were clearly shown to induce the expression of proteins in the 67 kDa – 97 kDa range, whereas it was not clearly observed in the chicken, newborn bovine and foetal calf serum.

Gel-based and gel-free proteomic analyses of the OMPs identified from these serum experiments are summarized in **Table 4-2**. Thirty-three OMPs were identified by growing the eight isolates in different sera (**Table 4-2**). Of these, three OMPs (OmpA, OmpH_1 and Omp 16/Pal) were found in all isolates and two proteins (OmpW and Lpp/Pcp) occurred in seven isolates (the exception being the ovine isolate PM982). TbpA was identified in both bovine isolates grown in newborn and foetal calf sera. Other iron receptors were identified in porcine and ovine isolates. HasR was identified in both porcine isolates. HgbA was identified in the porcine isolate PM734 and the ovine isolate PM966. PfhR, HemR and PM0803 were identified only the porcine isolate PM684. None of these iron receptors were identified in the avian isolates. This could indicate that there were sufficient iron sources available for the growth of bacteria in chicken serum compared to pig and sheep sera. When these isolates were grown in BHIB, TadD and RcpA were found only in the avian and ovine isolates, but both of these proteins were

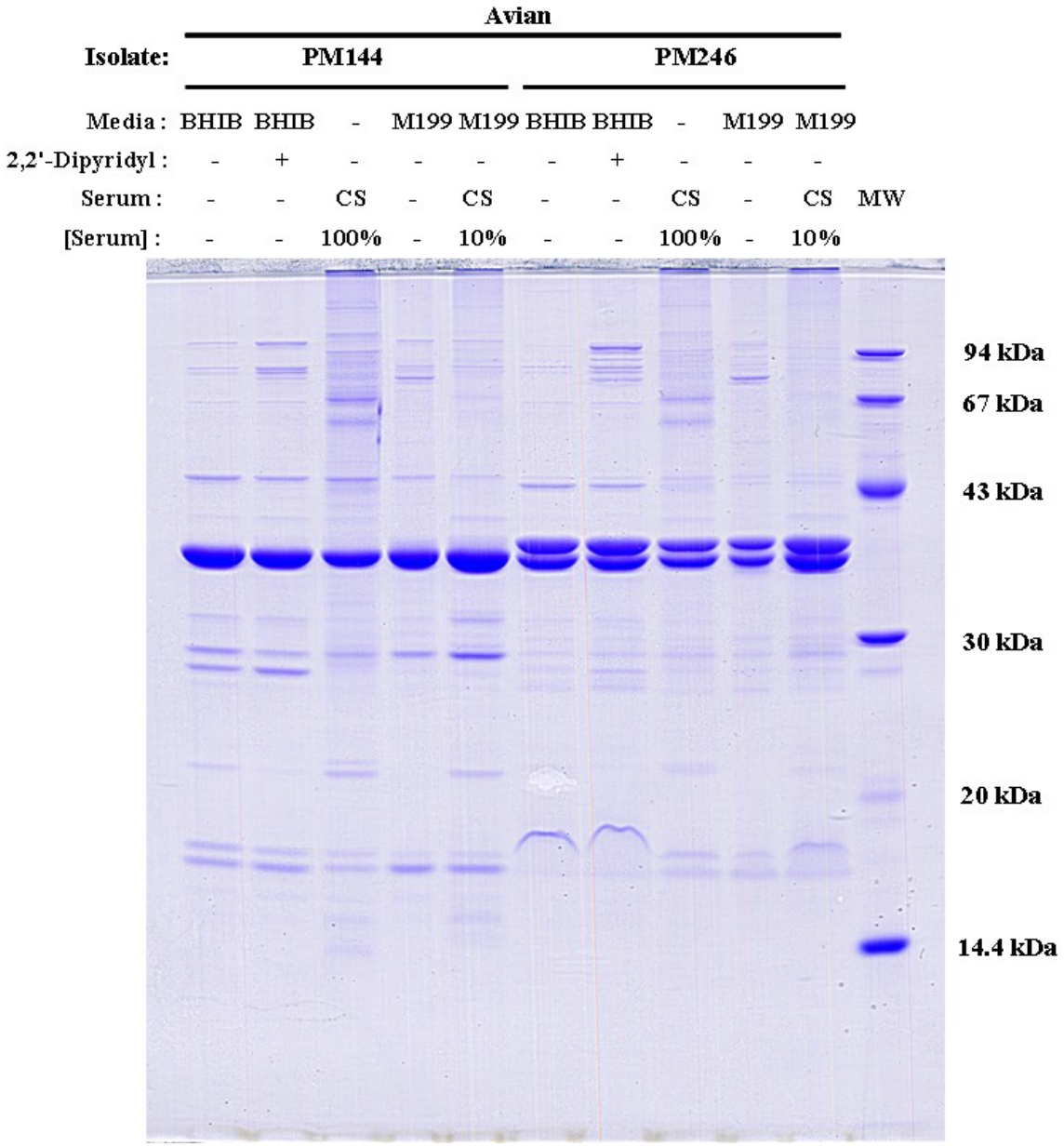


Figure 4-13. Comparison of the OMP profiles of the avian isolates PM144 and PM246 grown in BHIB, BHIB supplemented with dipyridyl, chicken serum, M199 and M199 supplemented 10% chicken serum. CS, chicken serum.

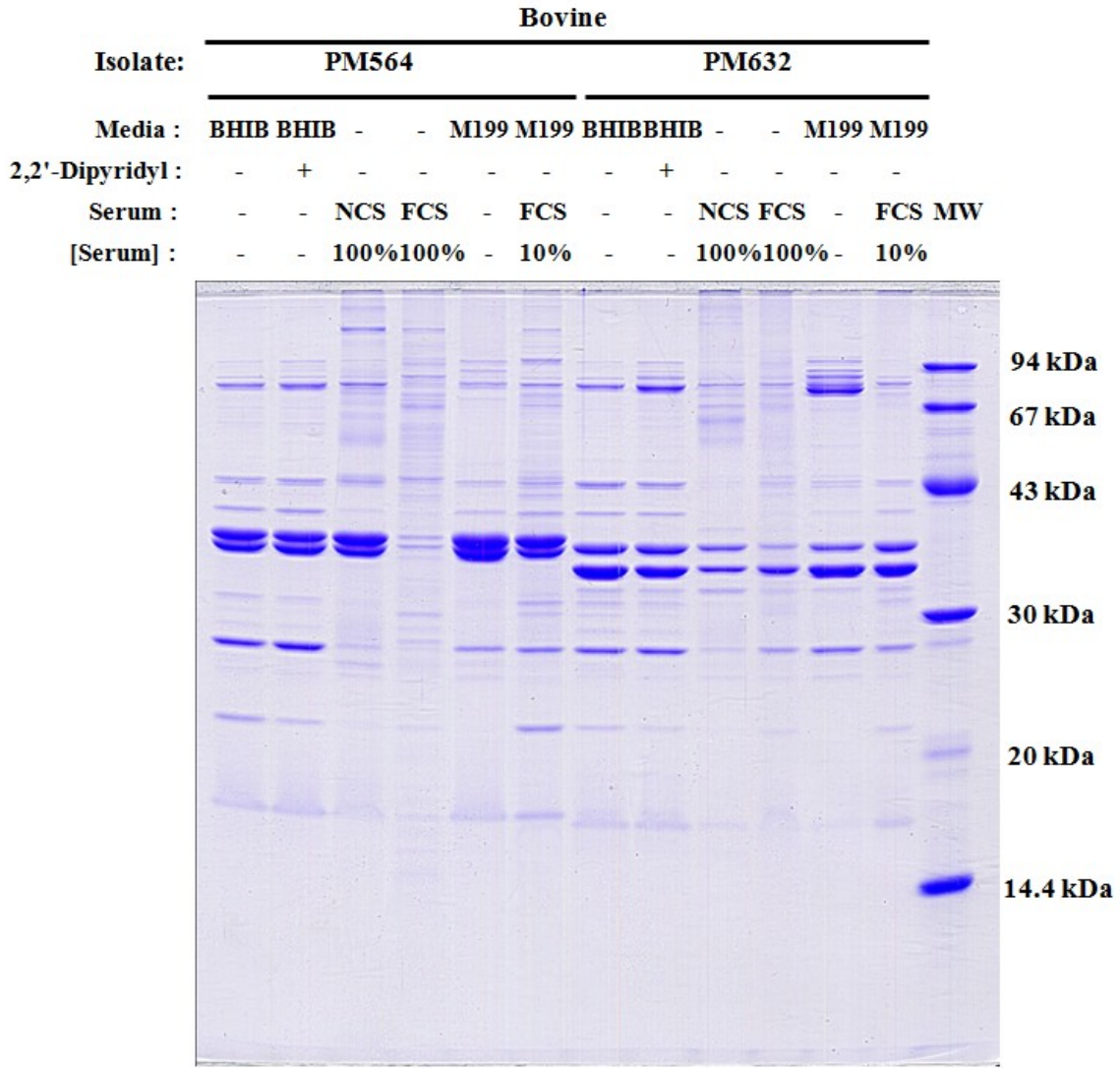


Figure 4-14. Comparison of the OMP profiles of the bovine isolates PM564 and PM632 grown in BHIB, BHIB supplemented with dipyridyl, newborn calf serum, foetal calf serum, M199 and M199 supplemented 10% foetal calf serum. NCS, newborn calf serum; FCS, foetal calf serum.

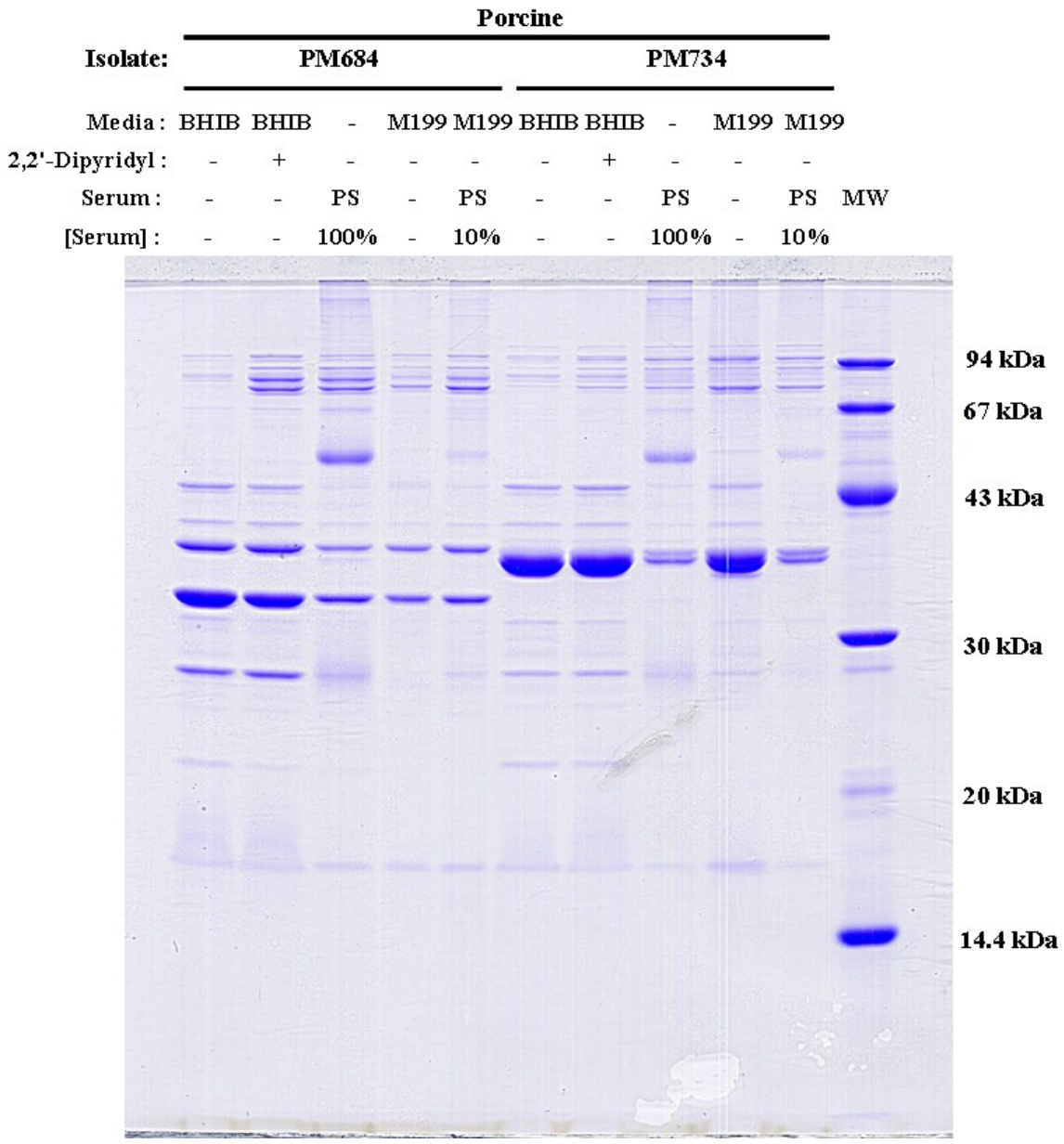


Figure 4-15. Comparison of the OMP profiles of the porcine isolates PM684 and PM734 grown in BHIB, BHIB supplemented with dipyridyl, pig serum, M199 and M199 supplemented 10% pig serum. PS, pig serum.

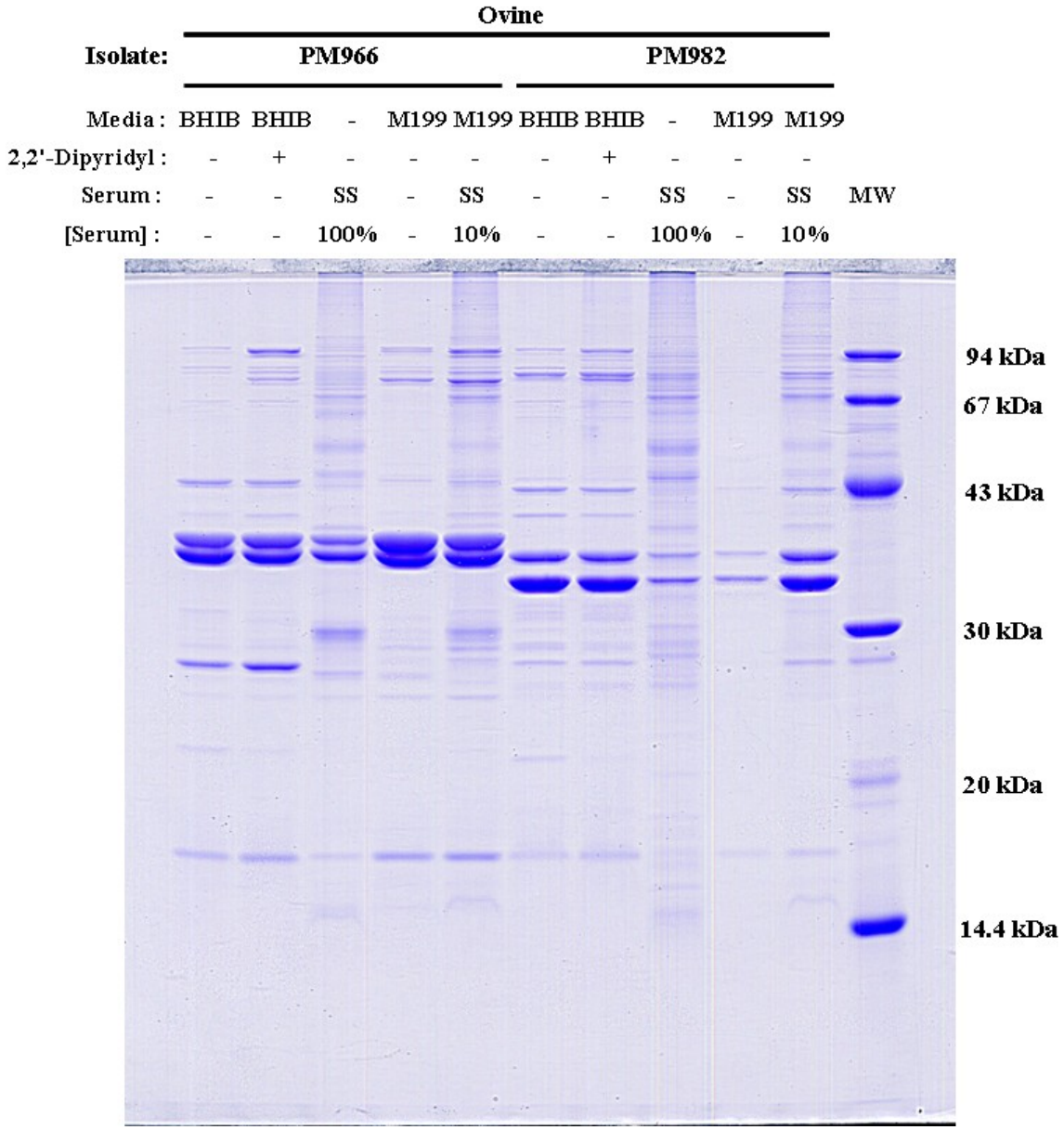


Figure 4-16. Comparison of the OMP profiles of the ovine isolates PM966 and PM982 grown in BHIB, BHIB supplemented with dipyridyl, sheep serum, M199 and M199 supplemented 10% sheep serum. SS, sheep serum.

identified in the bovine isolate PM564 after growth in serum. Similar to growing the avian isolates in BHIB, Opa was identified only in the avian isolates when grown in chicken serum. Proteins from serum (complement, haemoglobin, immunoglobulins, apolipoprotein, serotransferrin, albumin, vitronectin and LPS-binding proteins) as well as proteins from other bacterial cell compartments such as the cytoplasm (e.g. ribosomal proteins and TufA), inner membrane (e.g. PntA, NqrA and AtpA) and periplasm (e.g. SurA) were also identified in this study (**Appendix Table 4-1**).

When the eight isolates of *P. multocida* were grown in M199 and M199 supplemented with serum, 36 OMPs were identified (**Table 4-3**). OmpA and Omp P6/Pal were identified in all isolates and OmpH_1 and FadL were identified in seven isolates. HexD was identified only in the bovine, porcine and ovine isolates. Compared to the growth in BHIB and in serum, more iron receptors (eight) were detected in this experiment. These included PfhR, HasR, HgbA, HemR, TbpA, HgbB and TonB-dependent iron receptors PM0741 and PM0803. The number of iron receptors identified after growth in M199 of six isolates was higher than the number identified after growth in M199 supplemented with serum associated with the animal host. The avian isolates expressed three iron receptors (PfhR, PM0741 and PM0803) in isolate PM144 and six iron receptors (PfhR, HasR, HgbA, HgbB, PM0741 and PM0803) in isolate PM246 when they were grown in M199. None of these iron receptors were identified when the cells were grown in M199 supplemented with chicken serum. For the bovine isolate PM564, five iron receptors (HasR, HgbA, HemR, TbpA and PM0741) were identified when the isolate was grown in M199 and three iron receptors (HasR, HgbA and TbpA) were identified when the isolate was grown in M199 supplemented with foetal calf serum. The other bovine isolate PM632 reduced six iron receptors when it was grown in M199 and only one receptor when it was grown in M199 supplemented with foetal calf serum. For the porcine isolate PM684, the same five iron

Table 4-2. Comparison of the OMPs identified from eight isolates of *P. multocida* grown in serum associated with their hosts of origin.

No	Protein name	Protein ID	Isolates from different animal hosts									
			Avian		Bovine				Porcine		Ovine	
			PM144	PM246	PM564		PM632		PM684	PM734	PM966	PM982
			CS	CS	FCS	NCS	FCS	NCS	PS	PS	SS	SS
1	OmpA	PM0786	+	+	+	+	+	+	+	+	+	+
2	OmpH_1	PM0388	+	+	+	+	+	+	+	+	+	+
3	Omp16/Pal	PM0966	+	+	+	+	+	+	+	+	+	+
4	OmpW	PM0331	+	+	+	+	+	+	+	+	+	-
5	Lpp/Pcp	PM0554	+	+	+	+	+	+	+	+	+	-
6	OmpH_3	PM0831	+	-	+	-	+	+	+	+	-	+
7	TbpA	-	-	-	+	+	+	+	-	-	-	-
8	HexD	PM0778	-	-	+	+	-	-	+	+	-	-

"+" indicates the presence of a protein, whereas "-" indicates that the protein is not observed. CS, chicken serum; FCS, foetal calf serum; NCS, newborn calf serum; PS, pig serum; and SS, sheep serum.

Table 4-2. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts									
			Avian		Bovine				Porcine		Ovine	
			PM144	PM246	PM564		PM632		PM684	PM734	PM966	PM982
			CS	CS	FCS	NCS	FCS	NCS	PS	PS	SS	SS
9	MetQ/PlpA	PM1730	-	-	+	+	+	+	-	-	-	-
10	MipA/OmpV	PM0998	-	-	+	+	+	+	-	-	-	-
11	OmpH_2	PM0389	+	+	-	-	-	-	+	-	+	-
12	RcpA	PM0852	-	+	+	+	-	-	-	-	-	-
13	HasR	PM1622	-	-	-	-	-	-	+	+	-	-
14	Opa	PM1025	+	+	-	-	-	-	-	-	-	-
15	PM0016	PM0016	+	-	-	-	-	-	+	-	-	-
16	Oma87	PM1992	+	-	-	-	-	-	-	-	+	-
17	PM1827	PM1827	-	+	-	-	-	-	-	-	+	-
18	Omp85 family protein YftM	PM1809	-	-	-	+	-	-	-	-	+	-
19	HgbA	PM0300	-	-	-	-	-	-	-	+	+	-
20	FadL	PM1069	-	-	-	-	+	-	-	-	-	-

Table 4-2. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts									
			Avian		Bovine				Porcine		Ovine	
			PM144	PM246	PM564		PM632		PM684	PM734	PM966	PM982
			CS	CS	FCS	NCS	FCS	NCS	PS	PS	SS	SS
21	PfhR	PM0040	-	-	-	-	-	-	+	-	-	-
22	PlpE	PM1514/PM1517	-	-	-	-	-	-	+	-	-	-
23	SmpA	PM1886	+	-	-	-	-	-	-	-	-	-
24	GlpQ	PM1444	-	+	-	-	-	-	-	-	-	-
25	Plp4	PM0586	-	+	-	-	-	-	-	-	-	-
26	RcpB	PM0851	-	+	-	-	-	-	-	-	-	-
27	LppC	PM0646	-	-	-	+	-	-	-	-	-	-
28	PM1826	PM1826	-	-	-	+	-	-	-	-	-	-
29	TadD	PM0846	-	-	-	+	-	-	-	-	-	-
30	VacJ	PM1501	-	-	-	+	-	-	-	-	-	-
31	HemR	PM0576	-	-	-	-	-	-	+	-	-	-
32	PM0803	PM0803	-	-	-	-	-	-	+	-	-	-
33	OstA	PM1600	-	-	-	-	-	-	-	-	+	-
Total number of proteins			11	12	11	15	10	9	14	9	11	4

receptors (PfhR, HasR, HgbA, HemR and PM0803) were identified after growth in both M199 and M199 supplemented with pig serum. The other porcine isolate PM734 expressed six iron receptors (PfhR, HasR, HgbA, HemR, PM0741 and PM1428) when it was grown in M199 and three receptors (PfhR, HasR and HemR) when it was grown in M199 supplemented with pig serum. For the ovine isolate PM966, two iron receptors (PfhR and HgbA) were identified when it was grown in M199 and one receptor (PfhR) was identified when grown in M199 supplemented with sheep serum. No iron receptors were identified in the ovine isolate PM982 when grown in M199 but this could be due to not having enough outer membrane materials for the analysis. When this isolate was grown in M199 supplemented with sheep serum, only one iron receptor (TbpA) was identified. These results indicated that growing the eight isolates in M199 was similar to growing them under iron-limited conditions because this culture medium induced the expression of a number of iron receptors (**Table 4-4**).

Considering the previously identified OMPs in **Chapter 3**, a trimeric autotransporter adhesin Hsf_2 was a new OMP that was identified only in the two avian isolates. Opa was previously identified only in the avian isolates grown in BHIB and chicken serum; it was also identified only in the avian isolates grown in M199 and M199 supplemented with chicken serum. PlpE was previously identified only in the avian isolates grown in BHIB, but it was also found in the bovine isolate PM632 grown in M199. NanH was another OMP which was previously identified only in the bovine isolates; it was also found only in bovine isolate PM632 when grown in M199 supplemented with foetal calf serum.

Table 4-3. Comparison of the OMPs identified from eight isolates of *P. multocida* grown in M199 and M199 supplemented with 10% serum associated with their hosts of origin.

No	Protein name	Protein ID	Isolates from different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			M199	M199+10%CS	M199	M199+10%CS	M199	M199+10%FCS	M199	M199+10%FCS	M199	M199+10%PS	M199	M199+10%PS	M199	M199+10%SS	M199	M199+10%SS
1	OmpA	PM0786	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2	OmpP6/Pal	PM0966	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3	OmpH_1	PM0388	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+
4	FadL/OmpP1	PM1069	-	-	+	+	+	+	+	+	+	+	+	-	+	-	+	+
5	OmpH_2	PM0389	-	+	+	+	+	+	-	-	+	-	+	+	+	+	-	-
6	OmpH_3	PM0831	+	-	-	-	-	-	+	+	+	+	+	-	+	+	+	+
7	Lpp/Pcp	PM0554	+	+	-	-	+	+	+	+	-	-	+	+	-	-	+	+
8	HexD	PM0778	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-	+

"+" indicates the presence of a protein, whereas "-" indicates that the protein is not observed. CS, chicken serum; FCS, foetal calf serum; PS, pig serum; SS, sheep serum.

Table 4-3. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			M199	M199+10%CS	M199	M199+10%CS	M199	M199+10%FCS	M199	M199+10%FCS	M199	M199+10%PS	M199	M199+10%PS	M199	M199+10%SS	M199	M199+10%SS
9	PfhR	PM0040	+	-	+	-	-	-	+	-	+	+	+	+	+	+	-	-
10	HasR	PM1622	-	-	+	-	+	+	+	-	+	+	+	+	-	-	-	-
11	HgbA	PM0300	-	-	+	-	+	+	+	-	+	+	+	-	+	-	-	-
12	Omp87/Oma87	PM1922	+	-	-	-	+	-	+	+	-	-	-	-	+	-	+	+
13	HemR	PM0576	-	-	-	-	+	-	+	-	+	+	+	+	-	-	-	-
14	TbpA	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	+
15	PM0741	PM0741	+	-	+	-	+	-	+	-	-	-	+	-	-	-	-	-
16	PM0803	PM0803	+	-	+	-	-	-	+	-	+	+	-	-	-	-	-	-
17	MipA/OmpV	PM0998	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-
18	OmpW	PM0331	-	+	-	+	-	+	-	+	-	-	-	-	-	-	-	-
19	Plp4	PM0586	-	-	+	+	-	+	-	-	-	-	-	-	+	-	-	-

Table 4-3. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			M199	M199+10%CS	M199	M199+10%CS	M199	M199+10%FCS	M199	M199+10%FCS	M199	M199+10%PS	M199	M199+10%PS	M199	M199+10%SS	M199	M199+10%SS
20	Opa	PM1025	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
21	GlpQ	PM1444	-	+	-	+	-	-	-	-	-	-	-	-	-	+	-	-
22	MetQ/PlpA	PM1730	-	-	-	-	+	+	-	-	-	-	-	-	-	-	+	-
23	Hsf_2	PM1570	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-
24	TolC	PM0527	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-
	Lipoprotein																	
25	E/OmpP4	PM1064	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-
26	NanH	PM0663	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-
27	PM1428	PM1428	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
28	PM1543	PM1543	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
29	PM1826	PM1826	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-

Table 4-3. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			M199	M199+10%CS	M199	M199+10%CS	M199	M199+10%FCS	M199	M199+10%FCS	M199	M199+10%PS	M199	M199+10%PS	M199	M199+10%SS	M199	M199+10%SS
30	RpcA	PM0852	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
31	TadD	PM0846	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
32	PM0016	PM0016	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
33	RlpB	PM1215	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
34	HgbB	PM0337	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
35	OstA	PM1600	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
36	PlpE	PM1514	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
	Total number of proteins		11	10	14	12	17	14	19	13	11	10	14	9	11	7	8	9

Table 4-4. Comparison of the outer membrane iron receptors identified from eight isolates of *P. multocida* grown under iron-limited conditions and in M199. "+" indicate the presence of that protein and "-" indicates that the protein was not observed.

No	Protein ID	Protein name	Isolates of different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199
1	PM0300	HgbA	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-
2	PM0576	HemR	+	-	+	-	+	+	+	+	+	+	+	+	+	-	+	-
3	PM1622	HasR	-	-	+	+	-	+	+	+	+	+	+	+	-	-	+	-
4	PM0803	TonB-dependent receptor	+	+	+	+	-	-	+	+	+	+	+	-	+	-	-	-
5	-	TbpA	-	-	-	-	+	+	+	+	+	-	-	-	-	-	+	-
6	PM0337	HgbB	+	-	+	+	-	-	-	-	-	-	+	-	+	-	+	-
7	PM1428	TonB-dependent receptor	-	-	-	-	-	-	-	-	+	-	+	+	+	-	+	-
8	PM0040	PfhR	+	+	+	+	-	-	-	+	+	+	+	+	+	+	+	-

Table 4-4. (Continued)

No	Protein ID	Protein name	Isolates of different animal hosts															
			Avian				Bovine				Porcine				Ovine			
			PM144		PM246		PM564		PM632		PM684		PM734		PM966		PM982	
			Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199	Iron-limited	M199
9	PM0336	HgbB	+	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-
10	PM0741		+	+	+	+	-	+	+	+	+	-	-	+	+	-	-	-
11	-	Hup	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
12	PM1282	TonB-dependent receptor	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-

4.3.5 OMPs profiles after growth on solid surfaces

Eight isolates of *P. multocida* were grown on different agar media: BHI agar, BHI agar supplemented with 5% sheep's blood, BHI agar supplemented with 5% sucrose and BHI agar supplemented with 5% sucrose and 0.08% Congo Red. The growth of the eight isolates on these agar media was observed. Four isolates (avian and ovine isolates) had non-mucoid colony types, whereas the other four isolates (bovine and porcine isolates) were mucoid. Growth on BHI agar, BHI agar supplemented with 5% sheep's blood, and BHI agar supplemented with 5% sucrose of all isolates was very similar. However, growth of six isolates on BHI agar supplemented with 5% sucrose and 0.08% Congo Red yielded black pigment which is usually an indicator of biofilm-forming ability. The bovine isolate PM632 produced a small amount of black pigment, whereas, interestingly, the ovine isolate PM982 could not grow on Congo Red agar media.

Preparation of outer membrane samples from isolates grown on agar plates was performed by scraping the cells from the plates and subjecting to the Sarkosyl extraction process. The isolates having non-mucoid colonies were easier to remove compared to those having mucoid colonies. The OMP profiles of the eight isolates grown on different agar media are illustrated in **Figures 4-17 to 4-20**. For the avian isolates (**Figure 4-17**), the OMP profiles resulting from growth on BHI agar, BHI blood agar and BHI agar supplemented with sucrose were very similar to that resulting from growth in BHIB. A few differences were observed including 55 kDa and 32 kDa bands (arrows 1 and 3 in **Figure 4-17**) which were expressed only after growth on Congo Red agar; higher expression of a 40 kDa band (HexD/GlpQ) (arrow 2 in **Figure 4-17**) after growth on BHI agar and BHI blood agar; higher expression of a 29 kDa band (arrow 4 in **Figure 4-17**) after growth on BHIB, BHI agar supplemented with sucrose and Congo Red agar; higher expression of a 28 kDa band (arrow 5 in **Figure 4-17**) after growth on the four agar media. The OMP profiles resulting

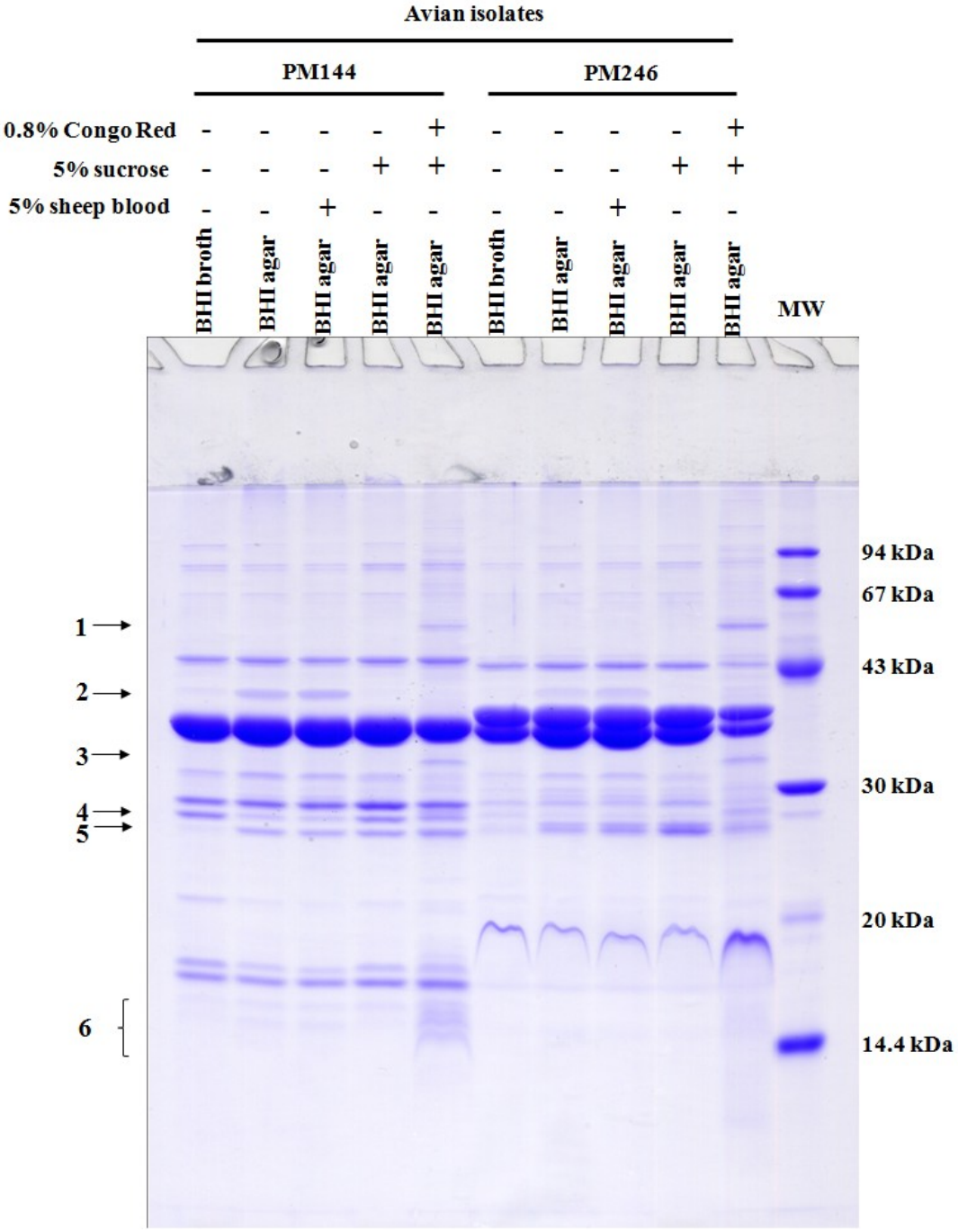


Figure 4-17. Comparison of the OMP profiles of avian isolates PM144 and PM246 grown on different agar media (BHI agar, BHI agar supplemented with 5% sheep's blood, BHI agar supplemented with 5% sucrose and BHI agar supplemented with 5% sucrose and 0.08% Congo Red) compared with the growth in BHIB. Numbers indicates differential expressed bands.

from growth on Congo Red agar were more complex with additional bands being expressed (arrows 1, 3 and 6 in **Figure 4-17**).

For the bovine isolates (**Figure 4-18**), similarly, the OMP profiles after growth on BHI agar, BHI blood agar and BHI agar supplemented with sucrose were similar to that of the growth in BHIB with a few quantitative changes (arrows 4-7 in **Figure 4-18**). 100 kDa, 55 kDa and 32 kDa bands (arrows 1, 2 and 3 in **Figure 4-18**) were expressed only after growth on Congo Red agar Whereas, the OMP profiles of cells grown on Congo Red agar were more complex, some of these additional bands were contaminant cytoplasmic, inner membrane and periplasmic proteins (**Appendix Table 4-3**). For the porcine isolates (**Figure 4-19**), the OMP profiles resulting from growth on BHI agar, BHI blood agar and BHI agar supplemented with sucrose were similar to that resulting from growth in BHIB with a few quantitative changes between 28 kDa - 30 kDa. The OMP profiles resulting from growth on Congo Red agar were complex and some of these additional bands were contaminant cytoplasmic, inner membrane and periplasmic proteins (**Appendix Table 4-3**). A number of additional bands (arrows 1 - 6 in **Figure 4-19**) were expressed only after growth on Congo Red agar. For the ovine isolate PM966 (**Figure 4-20**), the OMP profiles resulting from growth on BHI agar and BHI blood agar were similar including the expression of 40 kDa bands (HexD/GlpQ) (arrow 3 in **Figure 4-20**) in comparison to the growth in BHIB. This was also similar to the observation in the avian isolates. Higher expression of a 32 kDa band (arrow 5 in **Figure 4-20**) was observed after growth on BHI agar and BHI blood agar. Higher expression of proteins in the 28 kDa – 30 kDa range (No 6 in **Figure 4-20**) was observed after growth on all four agar media compared to the growth in BHIB. A number of additional bands (arrows 1, 2, 4 and 7 in **Figure 4-20**) were expressed only after growth on Congo Red agar. For ovine isolate PM982 (**Figure 4-20**), the OMP profiles resulting from growth on BHI agar, BHI blood agar and BHI agar supplemented with sucrose were similar to that oresulting from growth in BHIB.

Quantitative changes were also found in the region between 28 kDa to 30 kDa (arrow 6 in **Figure 4-20**). No OMP profile was obtained from the growth of this isolate on Congo Red agar because this isolate was unable to grow on this type of agar medium.

Gel-based and gel-free proteomic analyses of the OMPs obtained from the growth of the eight isolates on four types of agar media are summarized in **Table 4-5**. The OMPs obtained after growth on Congo Red agar were successfully analyzed by the gel-based method, but failed by the gel-free method because of the contamination of carried-over Congo Red to LC-MS/MS analysis. Forty-five OMPs were identified from the growth of these eight isolates on the four types of agar media. Six OMPs (HbpA, Hsf, hypothetical protein PM1543, SrfB, Skp and Wza) were newly identified with respect to those identified after growth in BHIB. HbpA is a TonB-dependent iron receptor which was filtered out by the prediction in **Chapter 2** and was identified by a literature search. Four OMPs (FadL, Oma87, OmpA and OmpH) were identified in all isolates and under all growth conditions. Most of the frequently identified OMPs (more than 16 times out of 32) were the core OMPs; these included FadL, Oma87, OmpA, OmpH, Pal, MipA/OmpV, HexD, PlpA/MetQ, Lipoprotein E/Omp P4, hypothetical protein PM1064, OmpW, TolC, Lpp, EstA and GlpQ (**Table 4-5**). Only HbpA was identified in all isolates after growth on Congo Red agar and in some isolates after growth on BHI agar. Six other iron receptors (TbpA, HgbA, PM1428, HasR, HmbR and PM0803) were identified in some isolates. TbpA was identified in both bovine isolates and one ovine isolate (PM966) in almost all types of agar media. Opa was identified only in the avian isolates after growth in all types of agar media, whereas NanH was identified only in the bovine isolate PM632 after growth on BHI agar, BHI blood agar and BHI agar supplemented with sucrose. TadD and RcpA were identified in two avian isolates, one bovine isolate (PM564), one porcine isolate (PM734) and one ovine isolate (PM966). These two proteins were identified only in the

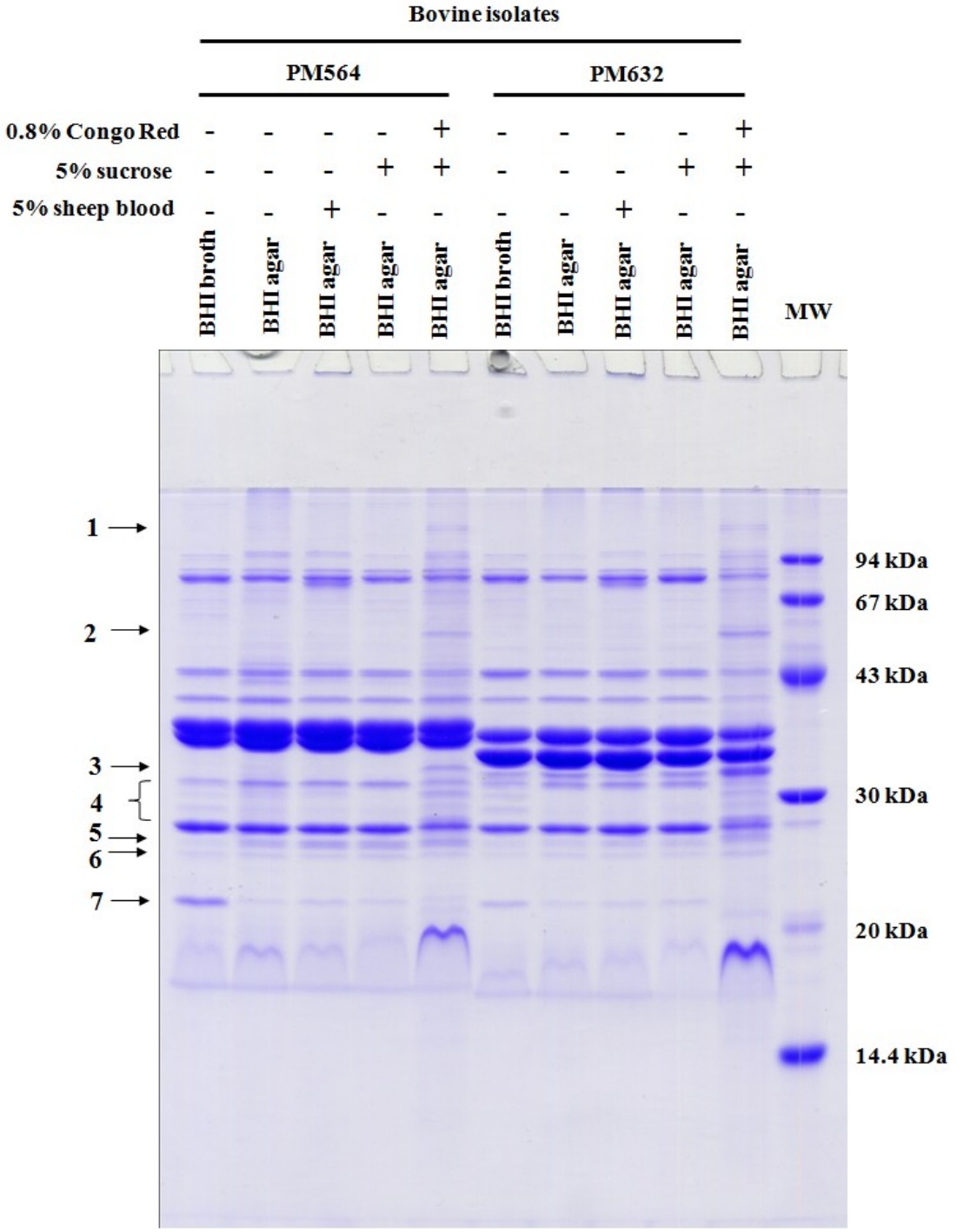


Figure 4-18. Comparison of the OMP profiles of bovine isolates PM564 and PM632 grown on different agar media (BHI agar, BHI agar supplemented with 5% sheep's blood, BHI agar supplemented with 5% sucrose and BHI agar supplemented with 5% sucrose and 0.08% Congo Red) compared with the growth in BHIB. Numbers indicates differential expressed bands.

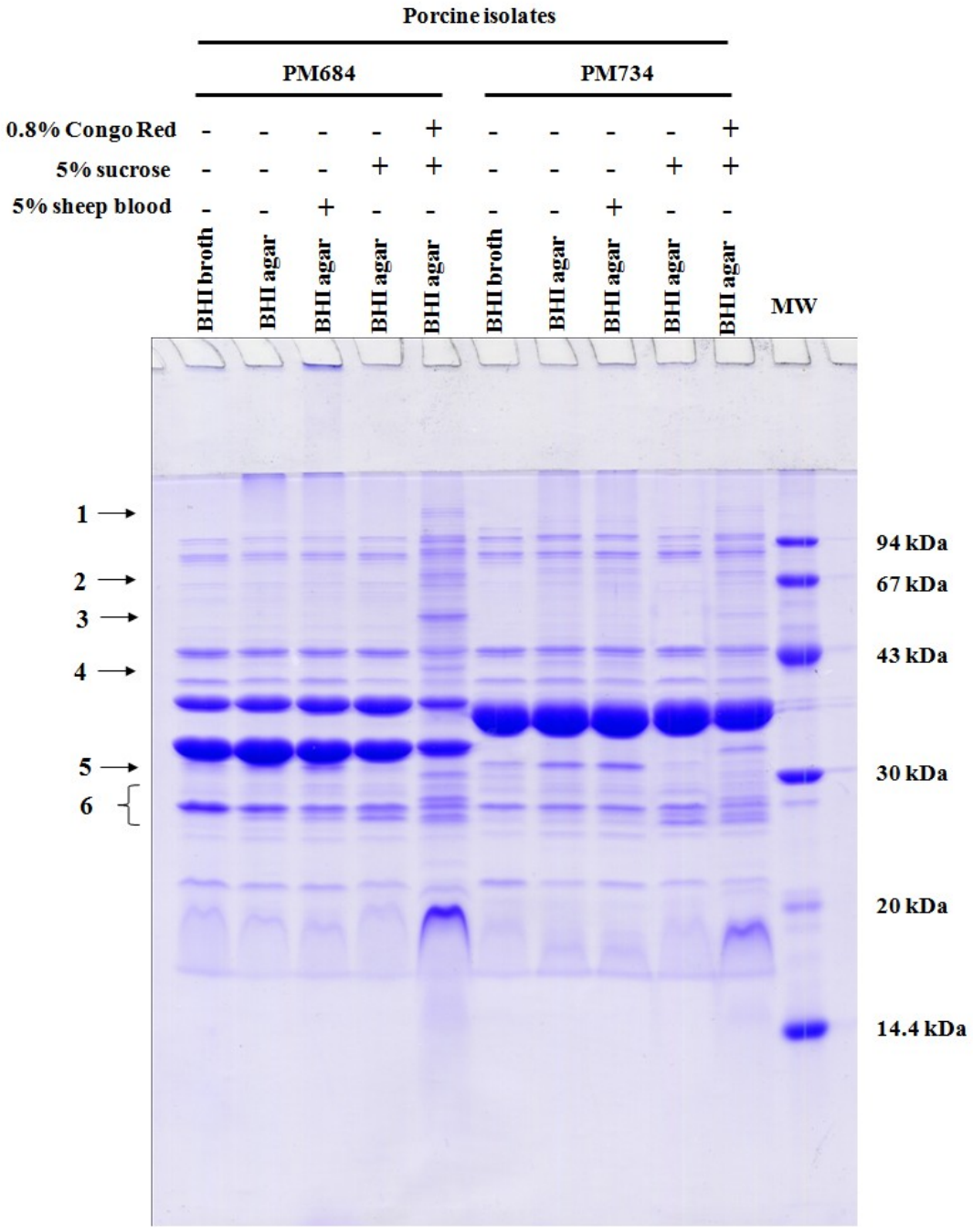


Figure 4-19. Comparison of the OMP profiles of porcine isolates PM684 and PM734 grown on different agar media (BHI agar, BHI agar supplemented with 5% sheep's blood, BHI agar supplemented with 5% sucrose and BHI agar supplemented with 5% sucrose and 0.08% Congo Red) compared with the growth in BHIB. Numbers indicates differential expressed bands.

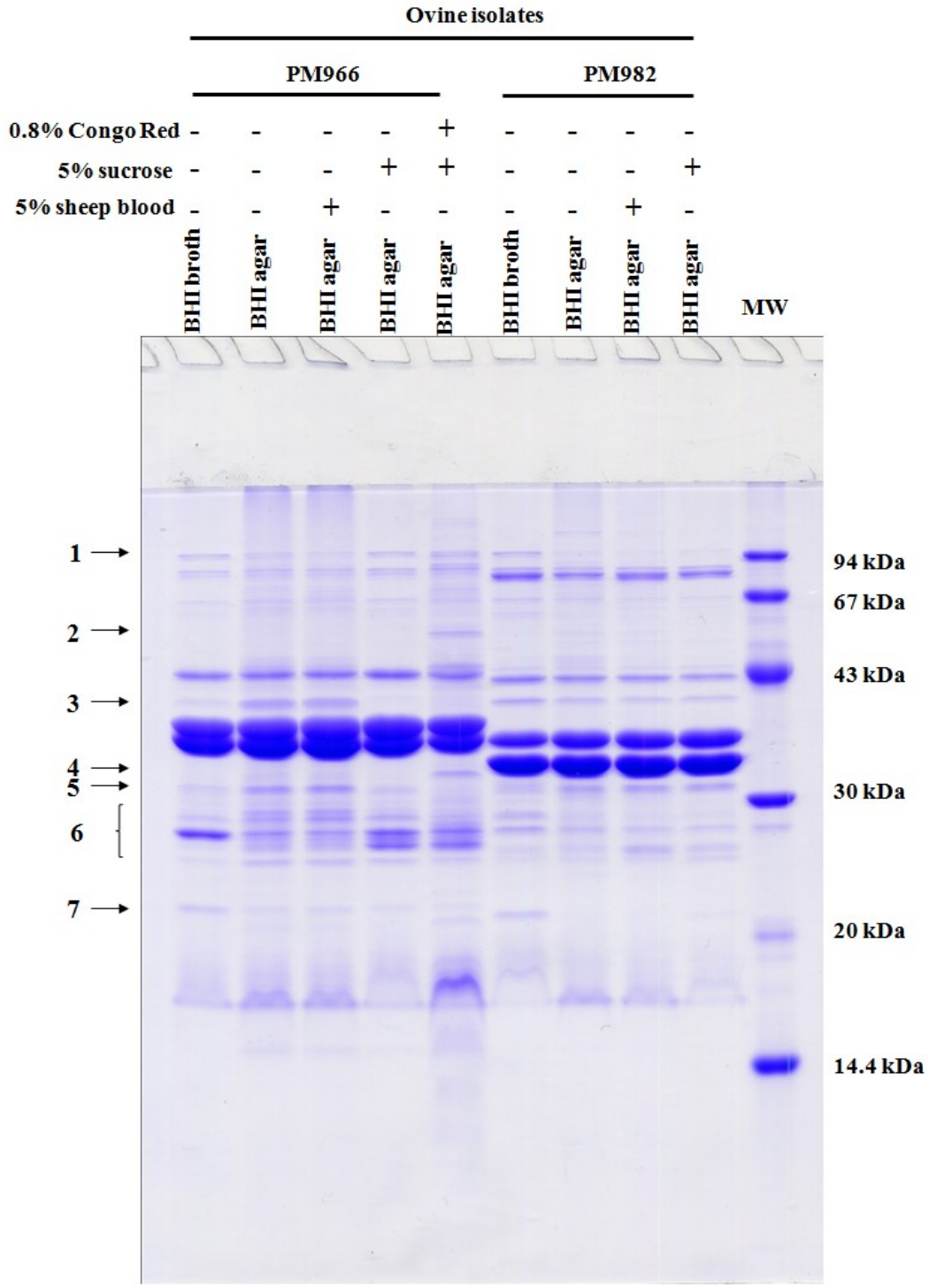


Figure 4-20. Comparison of the OMP profiles of ovine isolates PM966 and PM982 grown on different agar media (BHI agar, BHI agar supplemented with 5% sheep's blood, BHI agar supplemented with 5% sucrose and BHI agar supplemented with 5% sucrose and 0.08% Congo Red) compared with the growth in BHIB. Numbers indicates differential expressed bands.

Table 4-5. Comparison of the OMPs identified from the eight isolates of *P. multocida* grown on four types of agar media: BHI agar (BHIA), BHI blood agar (BHIA BLO), Congo Red agar (BHIA CR) and BHI agar supplemented with sucrose (BHIA SU).

No	Protein name	Protein ID	Isolates from different animal hosts																															
			Avian								Bovine								Porcine								Ovine									
			PM144				PM246				PM564				PM632				PM684				PM734				PM966				PM982			
			BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU
1	FadL	PM1069	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2	Oma87	PM1992	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3	OmpA	PM0786	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
4	OmpH	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
5	Pal	PM0996	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
6	MipA/ OmpV	PM0998	+	+	+	+	-	+	-	+	+	+	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+
7	HexD		-	+	-	+	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	
8	PlpA/MetQ	PM1730	+	-	+	+	+	+	-	+	-	-	+	-	+	-	+	+	+	+	+	+	+	-	+	-	+	-	+	+	+	+	+	+

"+" indicates the presence of a protein and "-" indicates that the protein is not observed.

Table 4-5. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts																															
			Avian								Bovine								Porcine								Ovine									
			PM144				PM246				PM564				PM632				PM684				PM734				PM966				PM982			
			BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU
9	Lipoprotein E/OmpP4	PM1064	+	+	-	+	+	+	-	-	+	+	+	+	+	+	-	+	+	+	-	+	+	+	-	-	+	-	-	-	+	+	+	+
10	PM0016	PM0016	+	+	-	+	+	+	-	+	+	+	-	+	-	+	-	+	+	+	-	+	-	+	-	+	+	+	+	+	-	+	-	+
11	OmpW	PM0331	-	+	+	+	+	+	-	-	+	+	-	+	-	+	-	+	+	+	-	+	+	+	+	+	-	+	+	-	-	-	-	+
12	TolC	PM0527	-	-	-	+	-	+	-	-	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+
13	Lpp	PM0554	+	-	+	+	-	+	-	+	-	-	-	+	+	+	+	+	+	-	-	-	+	+	-	+	+	-	+	+	-	-	+	+
14	EstA	PM0076	+	+	+	+	+	+	-	+	-	+	+	+	-	-	-	-	-	-	-	+	+	-	-	+	+	+	-	+	-	+	-	+
15	GlpQ	PM1444	+	+	-	+	+	+	-	-	+	+	-	+	+	+	-	-	+	+	-	-	+	+	-	-	+	+	-	-	-	+	-	-
16	OstA	PM1600	+	+	-	+	+	+	-	-	+	-	-	+	-	+	-	+	+	-	-	+	+	-	-	-	+	+	-	-	-	-	-	+
17	Plp4	PM0586	+	+	-	+	+	+	+	+	-	-	+	-	-	-	+	-	-	-	+	-	-	-	+	-	+	+	-	+	-	-	-	+
18	VacJ	PM1501	+	-	-	+	+	-	-	+	-	-	+	+	-	-	-	-	-	+	+	+	+	-	+	+	-	-	-	+	-	+	-	-

Table 4-5. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts																															
			Avian								Bovine								Porcine								Ovine									
			PM144				PM246				PM564				PM632				PM684				PM734				PM966				PM982			
			BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU
19	NlpB	PM1050	-	-	-	+	-	-	-	+	-	+	+	+	-	-	-	+	-	+	-	+	-	-	-	+	-	-	-	+	-	+	-	+
20	RcpA		+	+	-	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-
21	TadD	PM0846	+	+	-	+	+	+	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	-	-	-	-
22	OmpLA	PM1426	-	-	-	+	-	-	-	-	-	-	+	-	+	-	+	-	-	+	+	-	+	-	+	-	+	-	-	-	-	-	+	+
23	TbpA		-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
24	Opa	PM1025	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25	HgbA	PM0300	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	+	-	-	+	-	-	-	-
26	Hsf		-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-
27	ComL	PM1720	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	+	-	-	-	+	-	+	-	-	-	-	-	-	-
28	LppC	PM0646	-	-	+	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-

Table 4-5. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts																															
			Avian								Bovine								Porcine								Ovine									
			PM144				PM246				PM564				PM632				PM684				PM734				PM966				PM982			
			BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU
29	NanH	PM0663	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30	Omp85 family protein YftM	PM1809	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
31	LppB/ NlpD	PM1614	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
32	PM1826	PM1826	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-
33	PM1428	PM1428	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-
34	RlpB	PM1215	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-
35	HasR	PM1622	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-

Table 4-5. (Continued)

No	Protein name	Protein ID	Isolates from different animal hosts																															
			Avian								Bovine								Porcine								Ovine									
			PM144				PM246				PM564				PM632				PM684				PM734				PM966				PM982			
			BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU	BHIA	BHIA BL	BHIA CR	BHIA SU
36	HmbR		-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
37	PM1543	PM1543	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-
38	PM0803	PM0803	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
39	SrfB	PM1819	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
40	Skp	PM1993	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-
41	PM0442	PM0442	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42	SmpA	PM1886	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
43	Wza	PM1016	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
44	LspB	PM0056	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
45	HbpA	PM0592	-	-	+	-	-	-	+	-	+	-	+	-	-	-	+	-	+	-	+	-	+	-	+	-	-	-	+	-	-	-	+	-
Total number of proteins			18	17	13	25	19	20	9	20	16	15	19	19	15	16	15	18	15	16	14	18	22	14	19	15	24	16	16	18	11	16	11	19

avian isolates when grown in BHIB. Another OMP, Hsf, was identified in the avian isolate PM246 and the ovine isolate PM966.

4.4 Discussion

4.4.1 OMP changes in response to different stages of growth

This study examined the changes in the outer membrane proteome of two isolates of *P. multocida* and found slight changes in protein expression levels of HgbA, TolC, FadL and OmpW. This finding was consistent with the study in *M. haemolytica* which showed slight changes in expression of 18 kDa and 40.5 kDa proteins between log and stationary phase, of 94 kDa protein (HgbA) during log phase and a 24 kDa protein (OmpW) during stationary phase (Davies *et al.*, 1992). Similar results were also revealed in *Yersinia pestis* (Darveau *et al.*, 1980). The higher expression of TolC when grown in log-phase was consistent with the study by Evans and Poole (1999) who showed that a TolC homologue in *P. aeruginosa* was expressed at a the higher level during the log phase when compared to the lag phase. The OMP preparations in this study were usually obtained from *P. multocida* grown to late-log phase. These results confirmed that harvesting cells grown at different growth stages had a slight effect on the outer membrane proteome. However, greater changes in protein expression during different stages of growth might be obvious for the cytoplasmic and inner membrane proteomes because their proteome sizes are larger than the outer membrane proteome.

4.4.2 OMP changes in response to different rates of aeration

The OMP profiles of two isolates of *P. multocida* after growth under different rates of aeration were compared. The two isolates respond differently under different conditions of aeration. Growth with high aeration provides increased amounts of oxygen gas dissolved into the growth media compared to growth with normal aeration or no aeration (anaerobic). A higher shaking speed will also allow better access to nutrients and dilution of waste

products. Neither isolate showed high expression of iron receptors after growth with high aeration. This suggests that the isolates were able to gain access to enough amounts of iron. During growth with no aeration, there is a limited amount of dissolved oxygen gas and the mixing of nutrients and waste products is reduced. Under these conditions, the bovine isolate PM632 up-regulated expression of iron receptor proteins; the porcine isolate PM684 also expressed these proteins in slightly higher amounts than during growth with normal aeration. These findings were consistent with the growth of *M. haemolytica* under various rates of aeration and CO_2 tensions (Davies *et al.*, 1992). The authors observed the decreased expression of the iron receptor proteins when increasing the rate of aeration. Similarly, Frasch *et al.* (1976) compared the growth and OMPs of different Group B meningococcal strains under low and high rates of aeration. These authors observed increased growth rates when growing the bacteria with higher rates of aeration and also noted changes in the OMP content. This study demonstrated that there was a relation between the expression of iron receptors and oxygen concentration. A study in *P. aeruginosa* observed increased biosynthesis of haem during growth under low oxygen concentration resulting in an increased requirement of iron uptakes; consequently, there was increased biosynthesis of siderophores and up-regulated expression of iron receptors (Cox, 1986). A possible explanation could be the requirement of sufficient oxygen for bacterial metabolism. Cytochrome is an example of an oxygen-binding haem protein which contains haem as its structural component (Leys *et al.*, 2000). If oxygen concentration decreases, this can induce up-regulation of expression of cytochrome which then stimulates an increase of haem biosynthesis.

4.4.3 OMP changes in response to iron limitation

Iron limitation is a growth condition that is similar to *in vivo* condition in which iron is scarcely available to bacteria. This study used a synthetic iron chelator, 2,2′-dipyridyl, to sequester available iron in the growth media thereby creating the iron-limted conditions.

This iron chelator is widely used. However, the use of different types of iron chelators including synthetic and natural iron chelators could result in different responses of the bacterial cells. This was examined by Davies *et al.* (1992) who compared the growth of *M. haemolytica* in iron-limited growth conditions created by different types of iron chelators including 2,2′-dipyridyl, EDDA, desferal, ovotransferrin and bovine transferrin. Identical patterns of the iron-regulated protein expression were found after growth in the presence of 2,2′-dipyridyl, EDDA and desferal, but slightly different patterns were observed after growth in the presence of ovotransferrin and bovine transferrin. The use of synthetic iron chelators may not be fully representative of the *in vivo* iron-limited condition compared to the use of natural chelators. This provides evidence that bacteria are able to adjust the expression of specific iron receptors in order to respond to different levels of iron limitation created by different types of chelators or to different types of available iron compounds.

In this study, eight isolates of *P. multocida* were grown under iron-limited condition created by the presence of 2,2′-dipyridyl. The growth rates of all isolates were reduced in comparison to the growth in the absence of the iron chelator. This was consistent with studies in other bacteria such as *S. choleraesuis* (Ho *et al.*, 2004) and *F. tularensis* (Lenco *et al.*, 2007). The growth of the eight isolates was obtimized by using different concentrations of 2,2′-dipyridyl, indicating their different tolerance levels to iron limitation. The growth of the porcine isolates was considerably inhibited by using the lowest concentration of dipyridyl (45 and 90 μM). A study in *Porphyromonas gingivalis* also observed that virulent strains were more tolerant to iron-limited conditions than non-virulent strains (Grenier *et al.*, 2001). Proteomic analyses of the OMP profiles obtained after growth under iron-limited conditions showed that these eight isolates expressed different numbers of iron receptors and expressed more iron receptors in comparison to the same bacterial cells grown under iron-replete conditions. HgbA was expressed in all

isolates under iron-replete and iron-limited conditions. The avian isolates expressed only HgbA under iron-replete conditions. However, isolate PM144 expressed seven proteins (HgbA, HemR, PM0803, two HgbB proteins, PfhR and PM0741) and isolate PM246 expressed eight proteins (HgbA, HemR, HasR, PM0803, two HgbB proteins, PfhR and PM0741) under iron-limited growth conditions. Only HasR was additionally identified in isolate PM246. Under the iron-replete conditions, two iron receptors (HgbA and TbpA) were identified in the bovine isolate PM564 and four proteins (HgbA, HasR, PM0803 and TbpA) were identified in the bovine isolate PM632. When these isolates were grown under iron-limited conditions, isolate PM564 additionally expressed HemR and isolate PM632 additionally expressed HemR and PM0741. The porcine isolate PM684 expressed four iron receptor proteins (HgbA, HemR, HasR and TbpA) under iron-replete conditions but expressed an additional five proteins (PM0803, PM1428, PfhR, PM0741 and PM1282) under iron-limited growth conditions. The other porcine isolate PM734 expressed seven iron receptors (HgbA, HemR, HasR, PM0803, two HgbB proteins and PM1428) under iron-replete condition and two additional proteins (PfhR and Hup) under iron-limited conditions. Three iron receptors were identified in the ovine isolates under iron-replete condition; HgbA, PM1428 and HgbB for isolate PM966 and HgbA, TbpA and HgbB for isolate PM982. Under iron limited conditions, isolate PM966 expressed additional five proteins (HemR, PM0803, HgbB, PfhR and PM0741) and isolate PM982 expressed additional four proteins (HemR, HasR, PM1428 and PfhR). These differences in the expression of the iron receptors under iron-limited growth conditions in the presence of 2,2´-dipyridyl showed strain-specific expression patterns. This was supported by the study of Klitgaard *et al.* (2010) who demonstrated that a siderophore receptor, CirA, was expressed in all tested strains under iron-limited conditions and three specific haptoglobin-haemoglobin receptors were expressed in moderately and highly virulent strains. These patterns could be due to adaptation of these isolates to various animal hosts having different types and availabilities of iron compounds. An explanation for the ability to

express multiple iron receptors in *P. multocida* was also provided by Bosch *et al.* (2004) who suggested that it allowed the bacteria to adapt to variation in the iron sources. Based on the prediction results in **Chapter 2**, the predicted outer membrane proteomes of the avian and porcine strains of *P. multocida* contain the same set of the iron receptor proteins. However, the avian and porcine isolates selected in this study expressed different sets of iron receptors under the same conditions. These results suggest that different regulatory mechanisms may exist in some strains which allow certain iron receptors to be expressed in some strains, but not others, from the same set of shared proteins. This proposed model is shown in **Figure 4-21**. The ancestral strains of *P. multocida* contain a set of genes encoding iron receptor proteins. When these strains were transferred to new host niches (Hosts A and B in **Figure 4-21**), they are confronted by various iron-limited growth conditions in different animal hosts. These strains require a mechanism which will trigger levels of iron compounds and will adjust expressions of appropriate iron receptors. However, the exact mechanism that controls the expression of different iron receptors in different strains of *P. multocida* remains unknown. Regulation at the gene transcriptionl level could be involved in this mechanism. Holmes *et al.* (2006) compared gene expression between a wild-type and ferric uptake regulator (Fur) mutant strains of *Campylobacter jejuni* grown in iron replete and iron-limited conditions. The authors found that the iron receptor genes were highly up-regulated in the wild-type and the mutant grown under iron-limited conditions, suggesting that a global regulatory protein, Fur is an important regulator protein which controls the expression of the iron receptors. Fur is known to regulate expressions of a number of genes, not only genes encoding iron receptors (Andrews *et al.*, 2003). Under iron-limited conditions, Fur is released from binding to an upstream position of iron-regulated genes and allows expression of these genes. Jackson *et al.* (2010) identified 300 Fur-repressed and 107 Fur-induced genes by using transcriptome analysis, in silico Fur box prediction and Fur titration. The majority of these genes encode unknown protein. Certain of them are involved in iron metabolism,

cell communication, intermediary metabolism and energy metabolism. Small non-coding RNAs are involved in post-transcriptional regulation and are known to base-pair to mRNAs and bind to proteins (Repoila & Darfeuille, 2009). These non-coding RNAs regulate translational efficiency and stability of mRNA and can modify protein activity. A small non-coding RNA NrrF was shown to be involved in Fur regulation of iron responsive genes in *N. meningitidis* (Metruccio *et al.*, 2009). Further study of the interaction between Fur and small non-coding RNA in the regulation of iron responsive genes in different isolates of P. multocida might be able to explain the model in **Figure 4-21**.

The results of this study agreed with the previous microarray study by Paustian *et al.* (2001) who compared differential gene expression in *P. multocida* strain Pm70 grown under iron-replete and iron-limited conditions by using 2,2′-dipyridyl These authors identified five iron receptors (HemR, HgbB/PM0336, HgbA, PM0803 and PM0741) that were up-regulated during growth under iron-limited conditions. Similar results were obtained in the present study. The same five proteins were identified in the avian isolates, but two additional proteins (HgbB/PM0337 and PfhR) were identified in isolate PM144 and three additional proteins (HasR, HgbB/PM0337 and PfhR) were identified in isolate PM246. This results comfirmed that isolates of *P. multocida* express various patterns of iron receptors in response to iron limitation.

4.4.4 OMP changes in response to growth in different animal sera

In this study, eight isolates of *P. multocida* were grown in sera from different animals including chicken serum, foetal calf serum, newborn calf serum, pig serum and sheep serum. Each isolate grew differently in these five sera. All isolates grew best in chicken and foetal calf sera compared to the other three sera. *P. multocida* isolates did not grow specifically better in serum associated with their hosts of origin. All isolates seemed to grow slower in newborn bovine, pig and sheep sera. A possible explanation for this could

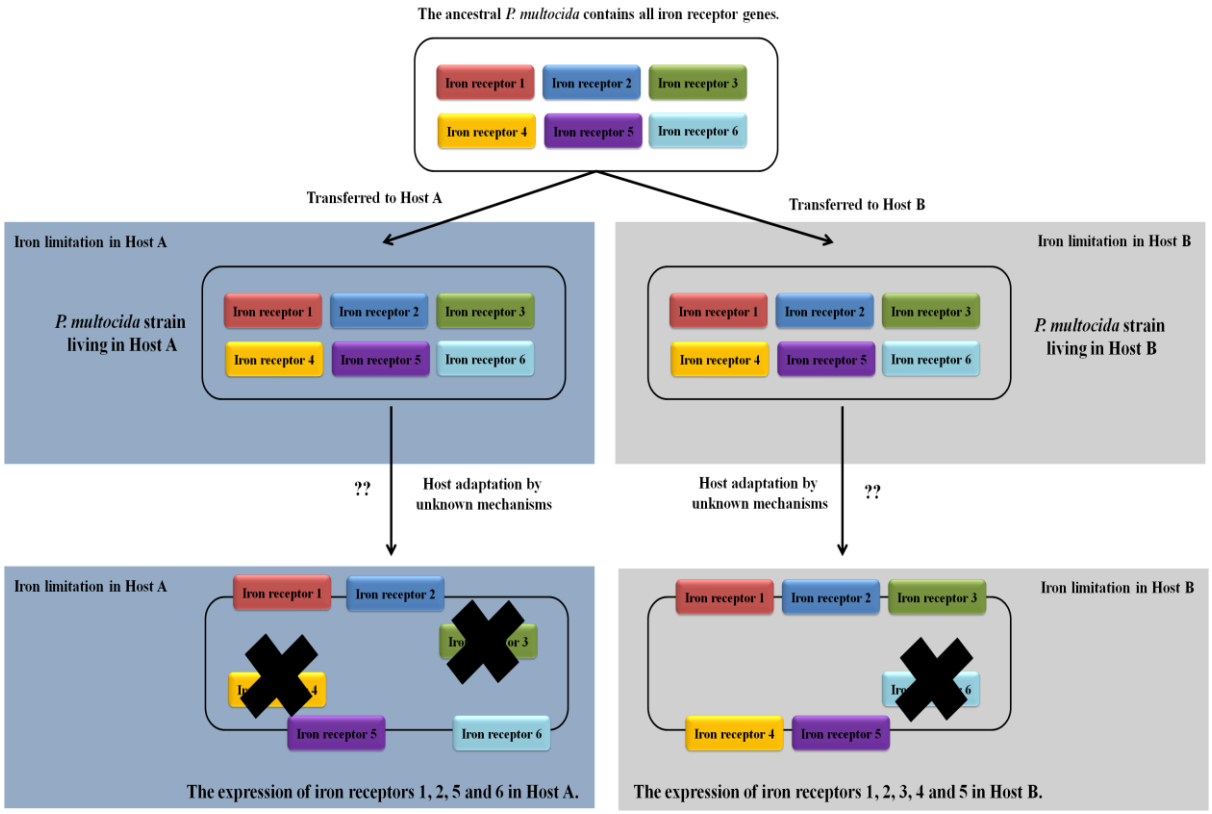


Figure 4-21. Proposed model to account for different expression of iron receptors in different isolates of *P. multocida* grown under iron-limited conditions. In this model, the ancestral *P. multocida* contains six iron receptor-encoding genes (1-6) labelled in different colors. When this ancestor has been transferred to different hosts (A and B), they may adapt to different host internal environments by changes in expression of iron receptor-encoding genes. The regulation of this adaptation process remains unclear.

be due to different serum compositions. A study by Miller *et al.* (2009) showed that pig serum contained a number of bacterial inhibiting proteins such as IgM, IgG, IgA and complement C3 and also provided potential iron sources such as transferrin, haemopexin and haptoglobin. The serum composition could vary in sera from different animal hosts of origin. Comparison of the OMP profiles obtained from the growth of the eight isolates in serum associated with their animal hosts was difficult, and more complex OMP profiles were obtained possibly due to the interference of serum proteins. However, by using a combination of proteomic methods, including the gel-based and gel-free techniques, certain OMPs could be identified from these samples. Frequently, the core OMPs were identified in all or some of the isolates. The reasons that the core OMPs were not identified in all isolates as shown in **Chapter 3** could be due to contamination with serum proteins which were highly abundant. These serum proteins could interfere with the Sarkosyl extraction process and also with the mass spectrometric process and analysis. This was confirmed by the identification of a number of serum proteins from the outer membrane samples. However, the patterns of certain host-restricted OMPs which were described in **Chapter 3** were also found in this study. For example, TbpA was associated only with bovine and ovine isolates and Opa was associated only with the avian isolates. This study did not observe increased expressions of iron receptors when isolates were grown in sera. This could be because there were available iron compounds within the serum. This study also observed changes in expression patterns of some host-restricted OMPs. RcpA and TadD were restricted to the avian isolates and the ovine isolate PM966 based on growth under iron-replete conditions, but were also identified in the bovine isolate PM564 after growth in serum. The reason for this observation could be that the genes in isolate PM564 were not expressed under iron-replete conditions, but were activated when grown in serum. Therefore, this study suggested that OMPs prepared from bacteria grown under a single growth condition may not represent all proteins within the

outer membrane proteome. Varying the growth conditions can improve the chances of identifying new or more OMPs.

The growth of two isolates of *P. multocida* in various combinations of growth media and serum showed that the addition of foetal calf serum into the media (BHIB and M199) did not increase the expression of the iron receptors of isolate PM632. However, the addition of pig serum into the media (M199) increased the expression of iron receptors in porcine isolate PM684. Comparison of the composition of BHIB and M199 revealed that BHIB was more nutritious for bacterial growth because it contains extracts from beef heart and calf brain supplemented with glucose and disodium hydrogenphosphate, whereas M199 contains salts, L-glutamine and sodium bicarbonate. Therefore, growing bacteria in BHIB is similar to growth under nutrient-rich conditions, whereas growing them in M199 is similar to growth under nutrient-limited conditions. The growth of the eight isolates of *P. multocida* in M199 and M199 supplemented with 10% serum associated with their animal hosts was further studied. The OMP profiles obtained from these growth conditions showed that growth in M199 induced the expression of iron receptors as clearly shown in isolates PM144, PM246 and PM632. When chicken and foetal calf sera were added to M199, the expression of these proteins declined and the growth rates of the bacteria increased, indicating that nutrients and iron conpounds present in the sera were available to the bacteria. However, the addition of pig and sheep sera did not result ina similar decrease in expression of iron receptors, indicating that nutrients and iron compounds were less available and these sera might contain some bacterial-inhibiting proteins such as immunoglobulins and complement. Proteomic analyses showed that growing these isolates in M199 induced the expression of iron receptors similar to growth under iron-limited condition, but in different patterns (**Table 4-6**). The reason for this could be due to different levels of iron restriction since M199 might contain very small

Table 4-6. Comparison of the iron receptor OMPs identified from the eight isolates of *P. multocida* grown under iron-replete and iron-limited conditions, and in M199 and M199 supplemented with 10% serum associated with their animal hosts.

No	Name of iron receptor	Protein ID	Isolates from different animal hosts																																	
			Avian								Bovine								Porcine								Ovine											
			PM144				PM246				PM564				PM632				PM684				PM734				PM966				PM982							
			Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum				
1	HgbA	PM0300	+	+	-	-	+	+	+	-	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	-	+	+	+	-	+	+	-	-				
2	HasR	PM1622	-	-	-	-	-	+	+	-	-	-	+	+	+	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	-	+	-	-				
3	HemR	PM0576	-	+	-	-	-	+	-	-	-	+	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	+	-	-	-	+	-	-				
4	PfhR	PM0040	-	+	+	-	-	+	+	-	-	-	-	-	-	-	+	-	-	+	+	+	-	+	+	+	-	+	+	+	-	+	-	-				
5	TbpA	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	+	+	-	+				
6	PM0803	PM0803	-	+	+	-	-	+	+	-	-	-	-	-	+	+	+	-	-	+	+	+	+	+	-	-	-	+	-	-	-	-	-	-				
7	PM0741	PM0741	-	+	+	-	-	+	+	-	-	-	+	-	-	+	+	-	-	+	-	-	-	-	+	-	-	+	-	-	-	-	-	-				
8	HgbB	PM0337	-	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	-	-	+	+	-	-				

"+" indicates the presence of the protein and "-" indicates the absence of the protein.

Table 4-6. (Continued)

No	Name of iron receptor	Protein ID	Isolates from different animal hosts																																			
			Avian								Bovine								Porcine								Ovine													
			PM144				PM246				PM564				PM632				PM684				PM734				PM966				PM982							
			Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum	Iron-replete	Iron-limited	M199	M199+serum				
9	PM1428	PM1428	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	+	-	+	+	-	-	-	+	-	-				
10	HgbB	PM0336	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+	+	-	-	-	-	-	-				
11	PM1282	PM1282	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
12	Hup		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-				
Total number of proteins			1	7	3	0	1	8	6	0	2	3	5	3	4	6	7	1	4	9	5	5	7	8	6	3	3	8	2	1	3	7	0	1				

amounts of available iron. Adding serum to M199 reduced the number of iron receptors expressed in almost all isolates.

4.4.5 OMP changes in response to growth on different solid surfaces

The OMPs of *P. multocida* are usually prepared after growth in broth (Boyce *et al.*, 2006; Davies *et al.*, 2003a, 2003b, 2003c, 2004) and no studies have compared the differences between growing this bacterium in broth and on a solid surface such as an agar plate. This study compared the growth of the eight isolates of *P. multocida* on four types of agar media, namely BHI agar, BHI agar supplemented with sheep's blood, BHI agar supplemented with sucrose and Congo Red and BHI agar supplemented with sucrose. Only growth on Congo Red agar resulted in the production of black pigment (indicative of biofilm formation) with the exception of ovine isolate PM982 which was completely inhibited on this type of agar plate. The Congo Red assay has been used as an indicator for biofilm formation in numerous studies (Abdallah *et al.*, 2009; Eroshenko *et al.*, 2010; Kimizuka *et al.*, 2009). With the exception of isolate PM982 which was unable to grow, all of the *P. multocida* isolates were able to form biofilms. The OMP profiles obtained from the four growth conditions were similar with a few quantitative changes. However, the OMP profiles obtained after growth on Congo Red agar plates were complex and additional protein bands were found to be contaminant proteins from other cell components. This could be due to the interference of Congo Red dye during Sarkosyl extraction and because of the carry-over of Congo Red in the final outer membrane fractions. These samples were unable to be analyzed by the gel-free method.

Proteomic identification of the OMP profiles obtained from these four growth conditions showed that the OMP profiles were similar to those after growth in broth. Most of the OMPs identified in almost all of the isolates and growth under the conditions represented the core OMPs described in **Chapter 3**. The reasons that the core proteins were not

identified in every isolate and under all growth conditions could be due to low level of expression and removal during Sarkosyl extraction. Surprisingly, a 60 kDa haemin-binding receptor (HbpA) was identified in all isolates after growth on Congo Red agar. HbpA is a TonB-dependent iron receptor protein which can bind to haemin and is negatively regulated by iron, manganese and haemin through the *fur*-independent pathway (Garrido *et al.*, 2003). This OMP was not identified when the isolates were grown under iron-limited conditions. Additionally, the nucleotide sequence of *hbpA* contains a hexanucleotide (AAAAAA) which can cause a programmed translational frameshift, resulting in different sizes of the encoding proteins (up to 40 kDa). This finding possibly explains the results observed from the gel-based analysis of the OMP profiles after growth on Congo Red agar. HbpA was identified in many gel bands that were excised after this growth condition.

The expression patterns of some host-restricted OMPs remained consistent with those after growth in broth. These included TbpA, Opa and NanH. Compared to the OMP profiles obtained after growth in broth, TadD and RcpA were additionally identified in the bovine isolate PM564 after growth on BHI agar and Congo Red agar, and in the porcine isolate PM734 after growth on Congo Red agar. Growing the eight isolates on these four types of agar did not cause stress from iron starvation because only one or two iron receptors were induced under these growth conditions. Under these growth conditions, a number of new OMPs were identified including Hsf, SrfB, Skp, Wza, and hypothetical protein PM1543. Interestingly, Hsf functions in adherence and this protein was identified only in the avian isolate PM246 and the ovine isolate PM982. Wza is acapsular polysaccharide transporter that was identified only in the avian isolate PM246 and only after growth on BHI agar. The other capsular polysaccharide transporter, HexD was identified in this isolate after growth in BHIB but this protein was absent when Wza was identified. Four adhesin OMPs were identified in this study and these included RcpA, TadD, Hsf and Opa. TadD and

RcpA were frequently found in the avian isolates and the ovine isolate PM966, while Hsf and Opa were identified only in the avian isolates and the ovine isolate PM966. These three isolates share the same non-mucoid colony type. Therefore, there might be a relationship between the expression of these adherence OMPs and the formation of non-mucoid colony morphology. Further characterization will be required to test this hypothesis.

4.4.6 Response of different isolates to the same growth conditions

In this chapter, the growth of different isolates of *P. multocida* under different growth conditions was examined. These growth conditions included iron limitation, growth in serum and media supplemented with serum, and growth on agar plates. Under the same growth condition, the eight isolates express the core OMPs, such as OmpA, OmpH and Oma87 which were identified in all isolates and under all growth conditions. The majority of these core OMPs function in biogenesis and integrity of the outer membrane and some of them (OmpA and OmpH) are highly abundant in the outer membrane, resulting in high chances of being identified. In addition to the core OMPs, certain host-restricted OMPs were strain-specific. For example, Opa was specific to avian isolates; TbpA was specific to bovine and ovine isolates; TadD and RcpA were specific to avian and ovine isolates; and NanH was specific to bovine isolates. Because these proteomic studies were based on eight selected representative isolates of *P. multocida*, the finding of host-specific or strain-specific OMPs was limited to a small numbers of isolates. The hypothesis that these host-restricted or strain-restricted proteins are likely to be involved in host adaptation could be confirmed by studying a large number of isolates of *P. multocida*.

4.4.7 Response of the same isolates to different growth conditions

When the same isolate of *P. multocida* was grown under different growth conditions, the changes in OMP profiles were quite obvious, especially in the comparison between iron-replete and iron-limited growth conditions as well as in the comparison between growth in M199 and M199 supplemented with serum. These results emphasize the fact that bacterial cells sense and respond to environmental changes by altering their outer membrane proteomes. For example, if the environment provides only trace amounts of iron, the bacterial cells will increase the expression of iron receptor OMPs to obtain sufficient iron for metabolism.

Chapter 5: General discussion and conclusion

5.1 Bioinformatic prediction of the outer membrane proteome of *P. multocida*

This study has used bioinformatic prediction approaches to predict individual members of the outer membrane proteome from two available *P. multocida* genome sequences. The use of consensus prediction by optimized criteria together with manual confirmation helped to narrow the list of predicted OMPs with a high degree of confidence. A number of studies applied consensus prediction and manual confirmation to their bioinformatic prediction of outer membrane proteomes and demonstrated improved degrees of confidence (Viratyosin *et al.*, 2008; Díaz-Mejía *et al.*, 2009; Heinz *et al.*, 2009). This prediction method can form the basis to understanding the outer membrane of this organism and will be useful for detailed characterization and functional studies of these proteins in the future. In the present study, the predicted outer membrane proteomes of *P. multocida* were validated by proteomic analyses of the Sarkosyl-extracted outer membrane proteomes. Detection of the predicted OMPs in the outer membrane samples (**Chapters 3 and 4**) comfirmed the confidence of this prediction.

However, the bioinformatic prediction workflow developed in this study could be further improved in a number of ways. New bioinformatic predictors are likely to be developed and the performance of these new predictors could improve the prediction. Selection criteria will be based on statistical parameters representing prediction performance of the program and capacity of the program for whole proteome prediction. Incorporating new predictors into the present prediction workflow will improve the performance of the present prediction method. At the consensus prediction stage, the selected criteria are highly stringent which efficiently removes most of the non-OMPs, but at the same time, these criteria also lead to the removal of a number of the true OMPs. This step significantly reduced the number of confidently predicted OMPs to a manageable and

reasonable size (4.9% of the avian strain genome and 4.7% of the porcine genome) compared to studies of other Gram-negative bacterial species (Berven *et al.*, 2006; Chung *et al.*, 2007). This study incorporated an additional step to analyze the filtered-out proteins and identify potential true OMPs. This step confirmed that there was a very low chance of losing true OMPs in this study (only 2%).

At the manual confirmation stage, various databases and bioinformatic tools were used to manually confirm the predicted proteins as being localized in the outer membrane. This step is time-consuming, particularly for larger bacterial genomes. In some cases, this study found that protein information deposited in the databases was incorrectly annotated, incomplete or unclear. Additional manual literature searchs together with sequence analyses (i.e., homology searchs and domain/motif searchs) was able to correct these errors. For example, the filamentous haemagglutinins PfhB_1 and PfhB_2 are very large proteins (234 and 232 kDa) which were predicted by some predictors to be OMPs. These two proteins were screened using the various criteria, but they were eventually confirmed to be secreted proteins during the manual confirmation step as a result of literature searchs for experimental evidence. Another example was PlpA/MetQ which was previously annotated in the UniProt protein database as PlpB. However, the annotation of this protein was later changed to PlpA/MetQ.

Some proteins, such as Ef-Tu (Kolberg *et al.*, 2008), can have multiple localizations under different growth conditions. Such proteins were unable to be predicted by the present bioinformatic study. Certain confidently predicted OMPs are classified as hypothetical proteins and there is little available information about the functions of these proteins. Experimental characterization of hypothetical proteins such as lipoproteins PM0554 and PM0016, which were identified in all isolates by proteomic methods, will be crucial for functional annotation and confirmation of localization. Functional characterization of these hypothetical proteins might be accomplished by attaching fluorescent expression tags

e.g., green fluorescent protein (GFP) to the OMP-encoding genes or proteins of interest which will allow visualization of the subcellular localization of those particular proteins once they are encoded (Phillips, 2001).

The bioinformatic studies could be further extended by a number of approaches. The principle of the bioinfomatic workflow could be applied to the prediction of the outer membrane proteome of other Gram-negative bacterial species. It could also be applied to predict other subcellular proteomes. Incorporating more computational programming steps, such as automatic data mining or word searching would reduce the time spent doing this manually. Systematic storage of predicted outer membrane proteomes and relevant information within a database or as an online web resource would be useful for further development of these results and for referencing. For example, *Echo*LOCATION is a database that provides subcellular localization of *E. coli* proteins obtained by bioinformatic prediction and experimental evidence (Horler *et al.*, 2009). The other database is CoBaltDB which stores prediction results of 43 subcellular localization-related predictors from 784 bacterial and archael proteomes (Goudenège *et al.*, 2010). The prediction of interactions between OMPs and proteins in other subcellular locations will provide the basis of understanding how the OMPs co-function with other proteins. This has been extensively studied in the OMPs involved in biogenesis and integrity of the outer membrane (Ruiz *et al.*, 2006) and in protein secretion pathways (Gerlach & Hensel, 2007). The prediction of protein-protein interaction might reveal novel molecular pathways. van Haagen *et al.* (2011) proposed a framework for prediction of protein-protein interaction in humans based on combination of different predictons and use of different data sources. These authors could identify novel protein-protein interaction pairs related to disease pathways. Applying this to *P. multocida* would be able to identify novel protein-protein interaction and assign function to hypothetical proteins.

The present study used avian and porcine strain genomes of *P. multocida* to predict the outer membrane proteomes. These two genomes do not contain the TbpA-encoding gene which is present in the bovine and ovine strains of this bacterium, shown in **Chapters 3 and 4**. This means that the predicted outer membrane proteomes of these two strains may not cover variations in the proteomes of the bovine and ovine strains of *P. multocida*. Since whole genome sequencing is now fast and less costly (Zhou *et al.*, 2010), comparative whole genomics of multiple isolates or pan-genomic analysis of *P. multocida* associated with diseases from different animal hosts will allow the *in silico* comparison of multiple *P. multocida* outer membrane proteomes and analysis of genome variation, different genomic structures and nucleotide polymorphism amoung different strains (Mira *et al.*, 2010).

Additional to the study of outer membrane proteomes, non-coding DNA regions could be investigated in order to understand regulatory mechanisms of the expression of the OMPs. Small non-conding RNA has been known as a regulator which allows a cell to adjust its physiology and metabolism in response to environmental changes (Repoila & Darfeuille, 2009). It can control stability and translational efficiency of mRNA and modulate protein activities including OMPs. A study by Song *et al.* (2008) demonstrated that a small non-coding RNA, *vrrA* in *Vibrio cholera*e could repress translation of *ompA* by base-pairing with the mRNA and *vrrA*-encoded protein VrrA induced formation of outer membrane vesicles. Identification of small non-coding RNA would allow better understanding of adaption of *P. multocida* isolates to different animal hosts.

5.2 The use of proteomic approaches to study the outer membrane proteome of *P. multocida*

5.2.1 Outer membrane extraction methods

Sarkosyl extraction was used to prepare the outer membrane-enriched fractions because this method yielded clean and reproducible OMP profiles (Hobb *et al.*, 2009). Sarkosyl

extraction was compared with extraction using other detergents as well as a spheroplasting method. Sarkosyl extraction was found to give the best representative and reproducible results. Because of the high reproducibility, this method has been used for OMP typing of *P. multocida* strains by Davies *et al.* (2003a, 2003b, 2003c, 2004). However, the use of differential solubilisation detergents might cause loss of certain loosely-associated outer membrane lipoproteins. Thus, complementing Sarkosyl extraction with other methods might reduce the possibility of losing some OMPs and would increase the confidence of OMP identification if the same OMPs were identified by different extraction methods. This idea was tested by comparing the OMPs identified from Sarkosyl extraction to cell envelope profiles obtained prior to solubilisation. Although the cell envelope profiles were complex and contained more proteins from the inner membrane and periplasm, additional outer membrane proteins were identified using this method compared to those from Sarkosyl extraction. Thus, a combination of these two approaches provides complementary data and increases the confidence of OMP identification. Other outer membrane extraction methods are also available including the use of sucrose density gradient centrifugation and could be used to compare with Sarkosyl extraction. For the future work, specific groups of OMPs could be extracted, such as surface-exposed OMPs, by surface-labelling proteomic techniques (Anaya *et al.*, 2007). A study in *Legionella pneumophila* used fluorescent labelling to identify surface-exposed OMP (Khemiri *et al.*, 2008a). This method could be further used to identify surface-exposed proteins from the outer membrane proteome of *P. multocida*.

5.2.2 Comparison of gel-based and gel-free proteomics

This study used two proteomic methods to prepare outer membrane protein peptides for mass spectrometric analyses: gel-based and gel-free methods. Combining these two methods improved the coverage and confidence of the OMP identification and have been used in a number of studies (Bridges *et al.*, 2008; Kouyianou *et al.*, 2010; Van Cutsem *et*

al., 2011). The study also used two gel-based methods. The first method involved cutting bands from the OMP profiles and the second involved analysing a single "band" containing the entire OMP sample by running the proteins for only a few millimeters into the resolving gel. The latter gel-based method reduced the numbers of bands to be cut and analysed by mass spectrometer which also meant a reduction on cost. However, bands containing the entire samples were more complex than those of the first method and could be a chance of losing some proteins. On the other hand, individual proteins could be identified and compared in multiple samples using the first gel-based method. These two gel-based methods complement each other. It was found that more transmembrane β-barrel OMPs were identified by the gel-based methods when compared to the gel-free method. Moreover, the proteins separated on the gel can be used for downstream analyses such as Western blotting and protein purification.

However, the gel-free method involves less steps and is less time-consuming compared to the gel-based methods. In this method, the OMPs are directly digested in solution. However, since the proteins are not denatured and remain embedded within the membrane, they might not be fully accessible to trypsin, resulting in a loss of certain proteins. The gel-free method was able to identify more outer membrane lipoproteins compared to the gel-based method. These lipoproteins are small and loosely associated with outer membrane. These proteins could be released from the membrane to the solution accessible for the enzyme. On the other hand, these small lipoproteins could be lost during preparation steps of the gel-based method. Therefore, these two methods complement each other. In some cases, such as OMPs obtained after growth on Congo Red agar plates, Congo Red had carried over into the OMP samples. In this case, it was not able to analyse the samples by the gel-free method and by the LC-MS. The gel-based method was the only method available for these samples. If the amount of proteins is very low, the gel-free method is the preferred option.

2D SDS-PAGE was also used in this study to separate the OMPs, of a single isolate of *P. multocida* (PM632). Proteins which were located close together, or within the same band, on the 1D gels could be clearly separated and quantitatively compared. The appropriate protocol for separation of the OMPs of *P. multocida* by 2D SDS-PAGE could be optimized and the method can be used to compare the OMP profiles from different isolates of *P. multocida*. 2D SDS-PAGE has more resolution compared to 1D SDS-PAGE. Protein bands which were close together could be separated. Whereas, more OMPs could be separated using 1D SDS-PAGE and multiple samples could be compared at the same time. The problem could be due to solubilisation of OMPs in rehydration buffer and extraction of OMPs from outer membrane could be difficult without using SDS as in 1D SDS-PAGE. ASB-14 in rehydration buffer of 2D SDS-PAGE could solubilise membrane proteins better CHAPS (Henningsen *et al.*, 2002) and it was used for separation of outer membrane proteome of an avian strain *P. multocida* (Boyce *et al.*, 2006). Boyce *et al.* (2006) identified less OMPs of *P. multocida* from 2D SDS-PAGE compared to from 1D SDS-PAGE.

Additionally, using different types of mass spectrometers can produce different results because they have varied levels of sensitivity, resolution and reproducibility. This study used MALDI-TOF/TOF MS/MS, LC-ESI/Q-TOF MS/MS and LC-Electron-transfer dissociation (ETD) ion trap MS/MS. MALDI-TOF/TOF MS/MS is limited to the sample complexity. Samples containing a large number of proteins will not be suitable for this mass spectrometer, but it is perfectly appropriate to analyze samples obtained from the excised gel bands. LC-MS/MS is suitable for the analysis of more complex samples but is sensitive to impurities of the samples such as salt contamination. Therefore, more proteins can be identified by using newer versions of the mass spectrometer with higher resolution and sensitivity.

5.3 Comparative outer membrane proteomics of eight disease isolates of *P. multocida* associated with different animal hosts grown under different growth conditions

This study compared the outer membrane proteomes of eight isolates of *P. multocida* associated with diseased chickens, cattle, pigs and sheep grown under different growth conditions. First, the outer membrane proteomes were characterized from growth of these isolates in iron-replete medium, BHIB, (**Chapter 3**) and these represented 52% of the predicted avian strain outer membrane proteome and 48% of the predicted porcine strain outer membrane proteome. Then, different growth conditions were examined to detect additional OMPs in these isolates (**Chapter 4**). The conditions included growth under iron-limited condition, growth in serum and in media supplemented with serum, and growth on agar plates as a potential biofilms condition. Additional OMPs were identified in these growth conditions which increased the total percentage to 62% of the predicted avian strain outer membrane proteome and 57% of the predicted porcine strain outer membrane proteome. This study confirms that all OMPs in the outer membrane proteome will not be expressed under a single growth condition and varying the growth conditions allows the detection of additional OMPs. These additional OMPs could be essential for adaptations of *P. multocida* to particular growth conditions.

The results in **Chapter 4** showed that the outer membrane proteome can be changed by varying degrees when the growth conditions are altered. Under the same growth conditions, the eight isolates responded by expression of the same core OMPs and variable numbers of minor OMPs. Some of these variable OMPs can be specific to an isolate or to a host, such as Opa, which was detected only in avian isolates regardless of the growth conditions. Under different growth conditions, each isolate can express the same core OMPs and variable numbers of minor OMPs which may be specific to the growth condition. For example, growth of these eight isolates in iron-limited conditions or in the medium M199 induced expression of more iron receptor OMPs compared to those grown

in iron-replete conditions (in BHIB). These iron receptors may help the bacteria to efficiently and sufficiently acquire iron from the host environment. Growth of these isolates on Congo Red agar induced the expression of an iron receptor HbpA which was not detected after growth under iron-limited or iron-replete conditions. This could provide evidence that *P. multocida* expresses different iron receptors in response to available iron sources. These iron receptors can be potential targets for the inhibition of *P. multocida* infection in different animal hosts. The remaining non-detected OMPs reported by the various predictions (**Chapter 2**) may be observed when growing these isolates in other growth conditions such as *in vivo* conditions.

Overall, the OMPs identified in this study (**Chapters 3 and 4**) were grouped into core and host-restricted OMPs, which will be discussed in **sections 5.3.1 and 5.3.2**. All OMPs from different isolates and growth conditions reported in **Chapters 3 and 4** were identified by at least one peptide with significant statistical scores. Using a combination of the gel-based and gel-free methods, can increase the confidence of the OMP identification and can be helpful to validate the identification of OMPs in the same sample. If the same OMP is detected multiple times in the sample analyzed by different methods, it will be confident to conclude the presence of this protein in the sample. Core OMPs are identified in all eight isolates; however examination of the expression of these OMPs in a greater number of isolates would confirm the results of this study. For the host-restricted OMPs, conclusions on host or growth condition specificities cannot be confidently made for a number of reasons. Some of these OMPs can be truly restricted to a host or growth condition, whereas the remaining proteins cannot be detected in all isolates or under all growth conditions because of loss during sample preparation and proteomic analyses or due to the low copy number of proteins expressed in the cell.

5.3.1 Core OMPs

Core OMPs were identified in all isolates under the same growth conditions. The Bam complex is involved in assembly and insertion of transmembrane β-barrel OMPs (Tommassen, 2010). The Bam complex consists of five proteins: BamA, BamB, BamC, BamD and BamE. Four of these proteins, BamA/Oma87, BamC/NlpB, BamD/ComL and BamE/SmpA, were expressed in all of the isolates. BamB was not identified in this study. The reason could be due to expression at undetectable level. Only BamA has been characterized in *P. multocida*; it is a surface-exposed protective antigen (Ruffolo & Adler, 1996). LptD/Imp/OstA is a member of the LptDE LPS assembly complex and was identified in all isolates. *In E.coli*, LptD forms a β-barrel structure in the outer membrane and has a periplasmic domain interacting with LptE (Freinkman *et al.*, 2011b). It has been proposed that LptE functions together with LptD, in the translocation of LPS to the outer leaflet of the outer membrane. However, LptE was not identified in the present study. Therefore, how the LptDE complex functions in *P. multocida* remains unknown. LolB was identified in this study although not in all of the isolates. These could be due to low level of expression or loss in Sarkosyl extraction. Theoretically, this protein should be considered as a core protein because it is part of the LolAB complex which is involved in the insertion of outer membrane lipoproteins into the outer membrane.

OmpA is highly abundant and provides structural integrity to the outer membrane by linking its C-terminal to the peptidoglycan layer (Carpenter *et al.*, 2007). The variation in molecular mass of this protein has been used to classify different OMP types of *P. multocida* (Davies *et al.*, 2003a, 2003b, 2003c, 2004). A previous study has shown that in *P. multocida* this protein functions in adherence to extracellular matrix proteins by using its extracellular loops (Dabo *et al.*, 2003). Pal and Lpp/Pcp are peptidoglycan-associated outer membrane lipoproteins, but how these proteins function in *P. multocida* has not been examined.

Certain enzymes function at the outer membrane in order to maintain the membrane integrity and decoration of the outer membrane for escaping from host defense mechanisms. OmpLA is a dimeric lipolytic enzyme in which the active site is at the outer part of the interface between two subunits. This phospholipase OmpLA is involved in the hydrolysis of phospholipids in the outer leaflet of the outer membrane to maintain the integrity of the outer membrane (Snijder & Dijkstra, 2000). GlpQ is a non-surface-exposed lipoprotein which removes glycerophosphocholine from the surface of host cells and changes glycerophosphocholine to choline which is a component of LPS in *H. influenzae* (Hatfaludi *et al.*, 2010). In *P. aeruginosa*, outer membrane esterase EstA functions in the hydrolysis of long chain acyl esters for use as a carbon source (Ohkawa *et al.*, 1979). OmpLA, GlpQ and EstA remain uncharacterized in *P. multocida*.

Different core transport OMPs were identified in this study. FadL is involved in the uptake of exogenous long chain fatty acids which can be utilized as carbon and energy sources (Zou *et al.*, 2008). The extracellular loops of this protein form a solvent-exposed hydrophobic groove which can bind to the long chain fatty acid. In *E. coli*, MetQ is part of the DL-methionine transport system (Merlin *et al.*, 2002). It is an outer membrane lipoprotein anchored at the inner side of the outer membrane. The protein binds to D-methionine and passes it into the cell via the ABC transporter MetD. Outer membrane porin OmpH is another highly abundant protein which has been used for OMP typing due to its molecular mass variation (Luo *et al.*, 1999; Davies *et al.*, 2003a, 2003b, 2003c, 2004). The protein is involved in the non-specific influx of various compounds (Pagès *et al.*, 2008). TolC is an energy-dependent multidrug efflux pump and a component of a type I secretion system (Hatfaludi *et al.*, 2008). These efflux pumps are involved in the multidrug resistance of the bacteria. A *tolC* mutant of *P. multocida* showed increased susceptibility to rifampin and acridine orange (Hatfaludi *et al.*, 2008). OmpW is a transporter for small hydrophobic molecules. It transports these molecules by a long and

narrow hydrophobic channel within the eight-stranded β-barrel (Hong *et al.*, 2006). The expression of OmpW was up-regulated under iron-limited conditions in *P. multocida* (Paustian *et al.*, 2001).

Another interesting core OMP is HexD which transports capsular polysaccharide across the outer membrane (Boyce *et al.*, 2000a). HexD was expressed in isolates associated with three capsular types (A, D and F). The capsular biosynthesis loci were only examined in isolates associated with capsular types A and B of *P. multocida* (Boyce *et al.*, 2000a). Major differences were observed in the genes involved in formation and assembly of the capsular polysaccharide (Boyce *et al.*, 2000a). Examination of these loci in other capsular types will allow us to understand the structural basis of the different capsular types. HexD is the only example of an OMP having a transmembrane α-helical structure. This study identified a second capsular polysaccharide transporter, Wza, in *P multocida*. This was expressed only in an avian isolate under certain growth conditions. This transmembrane α-helical protein is an octamer forming an elongated cylindrical structure with an interanal hydrophilic channel for accommodating the capsular polysaccharide (Collins & Derrick, 2007). This finding raises important questions about the presence of two capsule biosynthesis loci expressing very similar proteins, HexD and Wza. However, HexD was identified in all of the isolates, whereas Wza was identified in only one isolate. It is possible that this protein is not functional in the other isolates or it may be expressed under different growth conditions.

Three TonB-dependent iron receptors were identified in all isolates under different growth conditions. HgbA was present in all of the isolates after growth under iron-replete and iron-limited conditions. HgbA has been characterized in *P. multocida* and shown to be a haemoglobin receptor (Bosch *et al.*, 2002). A *hgbA* mutant did not affect the growth of *P. multocida*, possibly due to alternative expression of other iron receptor proteins. HemR was present in all of the isolates after growth under iron-limited conditions. In *Yersinia*

enterocolitica, HemR-expressing strains were able to utilize haem, haemoglobin, haptoglobin-haemoglobin, myoglobin, haemopexin and catalase as their iron sources (Bracken *et al.*, 1999). The haemin receptor, HbpA, was identified in all isolates after growth on Congo Red agar. This protein was previously identified in *P. multocida* after growth under iron-limited conditions and its expression is *fur*-independent (Garrido *et al.*, 2003). The relationship between HbpA expression and growth on Congo Red agar could be further examined.

There remain a number of core OMPs which have not been functionally characterized and could be targets for further studied. Based on the eight selected isolates of *P. multocida*, this study has identified a set of core OMPs which could be further analyzed and potential roles confirmed by examining larger numbers of isolates. Because these proteins are present in all isolates, they represent excellent targets for mutational studies. Strains of reduced virulence obtained by such mutational studies would be good candidates for drug and vaccine development. For example, an OmpA-like mutant in *Leptospira interrogans* was attenuated in virulence in the guinea pig and hamster models (Figueira *et al.*, 2007). By examining larger numbers of *P. multocida* isolates, molecular mass and sequence variation of certain core OMPs such as OmpA, OmpH and HgbA would reveal their adaptation and evolutionary histories (Davies *et al.*, 2001; Davies & Lee, 2004b). These could be investigated by sequencing the genes encoding these proteins.

5.3.2 Host-restricted outer membrane proteins

In addition to the core OMPs, this study identified host-restricted OMPs which were found in isolates associated with certain animal hosts and under certain growth conditions. Comparison of the OMP profiles prepared from the isolates associated with different hosts was inconclusive. These host-restricted OMPs could be truly associated with certain isolates or they could simply have been lost during OMP extraction in some isolates. Some proteins were not expressed in certain isolates grown under one set of growth

conditions but they were expressed in the same isolates grown under different growth conditions. For example, TadD and RcpA were frequently identified in the avian isolates and in ovine isolate PM966, but these proteins were also identified in bovine and porcine isolates after growth in M199 and on agar plates. TadD, RcpA and RcpB are members of the Tad locus which is responsible for the assembly of Flp pili and biofilm formation and this locus has not been characterized in *P. multocida*. TadD is essential for the assembly and function of the RcpABC proteins in *Aggregatibacter actinomycetemcomitans* (Clock *et al.*, 2008). Mutation in *rcpB* within this locus of *A. actinomycetemcomitans* resulted in an inability to form biofilms (Perez *et al.*, 2006).

Certain of the host-restricted OMPs were identified in the same isolates or in isolates from the same animal hosts under nearly all growth conditions. This provides indirect evidence to suggest that these proteins are specific to those isolates or animal hosts. TbpA is a transferrin binding protein which was frequently identified in the bovine and ovine isolates, and occasionally in the porcine isolate PM684. The TbpA protein of *P. multocida* is unusual because it does not require the presence of the outer membrane lipoprotein TbpB (Shivachandra *et al.*, 2005). The protein lacks one large extracellular loop which is possibly used to interact with TbpB (Ogunnariwo & Schryvers, 2001). The mechanism of transferrin binding and transport of TbpA in *P. multocida*, compared to that of TbpA in other species, could be further examined by sequencing and comparing different genes and the encoded proteins.

The opacity protein Opa is an OMP which was identified only in the avian isolates but under all growth conditions. This protein has been extensively studied in *Neisseria* species, but very little is known about this protein in *P. multocida*, for example why it is only expressed in the avian isolates. In *Neisseria*, Opa is a small surface-exposed transmembrane β-barrel protein with four extracellular loops (Dehio *et al.*, 1998). This

protein is involved in adherence and colonization of *Neisseria* sp. to host cells (Virji *et al.*, 1993). This protein could have a similar function in *P. multocida*.

A trimeric autotransporter adhesin Hsf is another adherence OMP that was identified in avian and ovine isolates. This protein has a lollipop-like structure and consists of three domains: the head, stalk and anchor domains (Linke *et al.*, 2006b). The anchor domain forms a transmembrane β-barrel structure in the outer membrane which functions in the autotransport of the head and stalk domains. The stalk domain has an extremely long. coiled coil-rich structure that stretches across the capsular layer. The head domain functions in attachment to host cells. The protein has been identified in many bacterial species, e.g. BadA in *Bartonella*, Hia and Hsf in *Haemophilus*, NadA in *Neisseria*, YadA in *Yersinia* and XadA in *Xanthomonas* (Linke *et al.*, 2006b). In *H. parasuis*, the trimeric autotransporter VtaA was expressed *in vivo* (Olvera *et al.*, 2010). In the present study, a trimeric autotransporter Hsf was expressed in avian isolates after growth in M199 and M199 supplemented with serum and in avian isolate PM246 and ovine isolate PM966 after growth on agar plates.

NanH is another OMP which was identified only in the bovine isolates. In *P. multocida*, this protein can cleave sialic acids from host cells allowing the bacteria to use them as nutrients (Sanchez *et al.*, 2004). Sialic acids can also be incoporated into their capsular polysaccharides or LPS to escape from host recognition (Hatfaludi *et al.*, 2010). NanH was also detected in persistently colonized rabbit strains of *P. multocida* (Sanchez *et al.*, 2004). It is possible that this protein will be expressed in other isolates when grown within the animal hosts.

The confirmation of these host-restricted proteins by proteomic methods, real-time PCR or Western blotting in a larger number of isolates will confirm the results of this study. Moreover, comparison of the OMP profiles obtained in this study with those obtained from

cells grown under *in vivo* growth conditions will allow a better understanding of the roles of OMPs in adaptation to different animal hosts. If these OMPs are true host-restricted proteins, they could be used as protein markers for particular strains/animal hosts of *P. multocida*.

In conclusion, this study used bioinformatic prediction to predict the outer membrane proteome of *P. multocida*. The outer membrane proteome was characterized, using a combination of proteomic methods, in eight representative isolates of *P. multocida* associated with different animal hosts and diseases. The isolates were grown under different growth conditions including in iron-replete and iron-limited media, in serum and media supplemented with serum, and on solid surfaces. Different isolates of *P. multocida* responded in different ways to the same growth condition. The same isolates altered their OMP profiles when the growth condition changed. The outer membrane proteome was devided into core and host-restricted OMPs. Core OMPs are present in, and important to, all isolates of this bacterium. Most function in biogenesis and integrity of the outer membrane, transport and receptor functions, and outer membrane associated enzymes. The host-restricted OMPs are present in certain isolates and probably have important roles in host adaptation.

Bibliography

Abascal, E. N., Guerra, A. C., Vázquez, A. S., Tenorio, V. R., Cruz, C. V., Zenteno, E., Contreras, G. P. & Pacheco, S. V. (2009). Identification of iron-acquisition proteins of *Avibacterium paragallinarum*. *Avian Pathol* **38**, 209-13.

Abdallah, F. B., Chaieb, K., Zmantar, T., Kallel, H. & Bakhrouf, A. (2009). Adherence assays and Slime production of *Vibrio alginolyticus* and *Vibrio parahaemolyticus*. *Braz J Microbiol* **40**, 394-398.

Abel, A., Sánchez, S., Arenas, J., Criado, M. T. & Ferreirós, C. M. (2007). Bioinformatic analysis of outer membrane proteome of *Neisseria meningitidis* and *Neisseria lactamica*. *Int Microbiol* **10**, 5-11.

Adler, B., Bulach, D., Chung, J., Doughty, S., Hunt, M., Rajakumar, K., Serrano, M., van Zanden, A., Zhang, Y. & Ruffolo, C. (1999). Candidate vaccine antigens and genes in *Pasteurella multocida*. *J Biotechnol* **73**, 83-90.

Adlam, C. & Rutters, J. M. (1989). *Pasteurella and Pasteurellosis.* UK: Academic Press Limited.

Aivaliotis, M., Corvey, C., Tsirogianni, I., Karas, M. & Tsiotis, G. (2004). Membrane proteome analysis of the green-sulfur bacterium *Chlorobium tepidum*. *Electrophoresis* **25**, 3468-74.

Al-Hasani, K., Boyce, J., McCarl, V. P., Bottomley, S., Wilkie, I. & Adler, B. (2007). Identification of novel immunogens in *Pasteurella multocida*. *Microb Cell Fact* **6**, 1-5.

Alteri, C. J. & Mobley, H. L. T. (2007). Quantitative profile of the uropathogenic *Escherichia coli* outer membrane proteome during growth in human urine. *Infect Immun* **75**, 2679-88.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-10.

Alves, J. R., Pereira, A. C. M., Souza, M. C., Costa, S. B., Pinto, A S., Mattos-Guaraldi, A. L., Hirata-Júnior, R., Rosa, A. C. P. & Asad, L. M. B. O. (2010). Iron-limited condition modulates biofilm formation and interaction with human epithelial cells of enteroaggregative *Escherichia coli* (EAEC). *J Appl Microbiol* **108**, 246-55.

Amory, H., Jean, D., Leveille, R., Higgins, R. & Vrins, A. (2006). *Pasteurella multocida* isolation in a horse with retropharyngeal infection. *J Equine Vet Sci* **26**, 364-369.

Anaya, C., Church, N. & Lewis, J. P. (2007). Detection and identification of bacterial cell surface proteins by fluorescent labeling. *Proteomics* **7**, 215-9.

Andrews, S. C., Robinson, A. K. & Rodriguez Quinones, F. (2003). Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**, 215-237.

Ang, C.-S., Veith, P. D., Dashper, S. G. & Reynolds, E. C. (2008). Application of 16O/18O reverse proteolytic labeling to determine the effect of biofilm culture on the cell envelope proteome of *Porphyromonas gingivalis* W50. *Proteomics* **8**, 1645-60.

Anwar, H., Brown, M. R., Britten, a Z. & Lambert, P. a. **(1983).** Disruptive effects of tris and sodium lauroyl sarcosinate on the outer membrane of *Pseudomonas cepacia* shown by fluorescent probes. *J Gen Microbiol* **129**, 2017-20.

Arkin, I. T. & Brunger, A. T. (1998). Statistical analysis of predicted transmembrane alpha-helices. *Biochim Biophys Acta* **1429**, 113-128.

Arun, S. J. & Krishnappa, G. (2004). Antigenic analysis of outer membrane proteins of biofilm and free cells of *Pasteurella multocida* A:1 in comparison with local isolate. *Int J Poultry Sci* **3**, 517-521.

van Alen, T., Claus, H., Zahedi, R. P., Groh, J., Blazyca, H., Lappann, M., Sickmann, A. & Vogel, U. (2010). Comparative proteomic analysis of biofilm and planktonic cells of *Neisseria meningitidis*. *Proteomics* **10**, 4512-21.

Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L. & Sankaran, K. (2006). A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol* **188**, 2761-2773.

Babu, M. M. & Sankaran, K. (2002). DOLOP--database of bacterial lipoproteins. *Bioinformatics* **18**, 641-643.

Babujee, L., Venkatesh, B., Yamazaki, A. & Tsuyumu, S. (2007). Proteomic analysis of the carbonate insoluble outer membrane fraction of the soft-rot pathogen *Dickeya dadantii* (syn. *Erwinia chrysanthemi*) strain 3937. *J Proteome Res* **6**, 62-9.

Bagos, P. G., Liakopoulos, T. D. & Hamodrakas, S. J. (2004a). Finding beta-barrel outer membrane proteins with a Markov Chain Model. *WSEAS Transactions on Biology and Biomedicine* **2**, 186-189.

Bagos, P. G., Liakopoulos, T. D., Spyropoulos, I. C. & Hamodrakas, S. J. (2004b). PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins. *Nucleic Acids Res* **32**, W400-4.

Bagos, P. G., Liakopoulos, T. D., Spyropoulos, I. C. & Hamodrakas, S. J. (2004c). A Hidden Markov Model method, capable of predicting and discriminating beta-barrel outer membrane proteins. *BMC bioinformatics* **5**, 29.

Bagos, P. G., Liakopoulos, T. D. & Hamodrakas, S. J. (2005). Evaluation of methods for predicting the topology of beta-barrel outer membrane proteins and a consensus prediction method. *BMC Bioinformatics* **6**, 1-13.

Baik, S.-chul, Kim, K.-mi, Song, S.-min, Kim, D.-su, Jun, J.-su, Lee, S.-gyu, Song, J.-young, Park, J.-uck, Kang, H.-lyun, & other authors. **(2004).** Proteomic analysis of the sarcosine-insoluble outer membrane fraction of *Helicobacter pylori* strain 26695. *J Bacteriol* **186**, 949-955.

Baltes, N. & Gerlach, G. F. (2004). Identification of genes transcribed by *Actinobacillus pleuropneumoniae* in necrotic porcine lung tissue by using selective capture of transcribed sequences. *Infect Immun* **72**, 6711-6716.

Bates, P. A., Kelley, L. A., Maccallum, R. M. & Sternberg, M. J. E. (2001). Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins* **5**, 39 - 46.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res* **28**, 235-42.

Berthiaume, F., Leblond, M.-F., Harel, J. & Mourez, M. (2010). Growth-phase-dependent expression of the operon coding for the glycosylated autotransporter adhesin AIDA-I of pathogenic *Escherichia coli*. *FEMS Microbiol Lett* **311**, 176-84.

Berven, F. S., Flikka, K., Jensen, H. B. & Eidhammer, I. (2004). BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. *Nucleic Acids Res* **32**, W394-9.

Berven, F. S., Karlsen, O. A., Straume, A. H., Flikka, K., Murrell, J. C., Fjellbirkeland, A., Lillehaug, J. R., Eidhammer, I. & Jensen, H. B. (2006). Analysing the outer membrane subproteome of *Methylococcus capsulatus* (Bath) using proteomics and novel biocomputing tools. *Arch Microbiol* **184**, 362-77.

Besemer, J. & Borodovsky, M. (2005). GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* **33**, W451-4.

Bhairi, S. M. (2001). *A guide to the properties and uses of detergents in biological system*. Germany: Calbiochem-Novachem-Corporation.

Bhasin, M., Garg, A. & Raghava, G. P. S. (2005). PSLpred: prediction of subcellular localization of bacterial proteins. *Bioinformatics* **21**, 2522-4.

Bigelow, H. & Rost, B. (2006). PROFtmb: a web server for predicting bacterial transmembrane beta barrel proteins. *Nucleic Acids Res* **34**, W186-8.

Birdsell, D. C. & Cota-Robles. **(1967).** Production and ultrastructure of lysozyme and ethylenediaminetetraacetate-lysozyme spheroplasts of *Escherichia colil*. *J Bacteriol* **93**, 427-437.

Blackall, P. J., Fegan, N., Chew, G. T. I. & Harnpson, D. J. (1998). Population structure and diversity of avian isolates of *Pasteurella rnultocida* from Australia. *Microbiol* **144**, 279-289.

Borrathybay, E., Sawada, T., Kataoka, Y., Ohtsu, N., Takagi, M., Nakamura, S. & Kawamoto, E. (2003). A 39kDa protein mediates adhesion of avian *Pasteurella multocida* to chicken embryo fibroblast cells. *Vet Microbiol* **97**, 229-243.

Bos, M. P., Robert, V. & Tommassen, J. (2007). Biogenesis of the gram-negative bacterial outer membrane. *Annu Rev Microbiol* **61**, 191-214.

Bosch, M., Garrido, M. E., Llagostera, M., Perez de Rozas, A. M., Badiola, I. & Barbe, J. (2002). Characterization of the *Pasteurella multocida hgbA* gene encoding a hemoglobin-binding protein. *Infect Immun* **70**, 5955-5964.

Bosch, M., Garrido, M. E., Pérez de Rozas, A. M., Badiola, I., Barbé, J. & Llagostera, M. (2004). *Pasteurella multocida* contains multiple immunogenic haemin- and haemoglobin-binding proteins. *Vet Microbiol* **99**, 103-12.

Bouhss, A., Trunkfield, A. E., Bugg, T. D. H. & Mengin-Lecreulx, D. (2008). The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev* **32**, 208-33.

Boyce, J. D., Chung, J. Y. & Adler, B. (2000). *Pasteurella multocida* capsule: composition, function and genetics. *J Biotechnol* **83**, 153-60.

Boyce, J. D., Chung, J. Y. & Adler, B. (2000b). Genetic organisation of the capsule biosynthetic locus of *Pasteurella multocida* M1404 (B:2). *Vet Microbiol* **72**, 121-34.

Boyce, J. D. & Adler, B. (2006). How does *Pasteurella multocida* respond to the host environment? *Curr Opin Microbiol* **9**, 117-22.

Boyce, J. D., Cullen, P. a & Adler, B. (2004). Genomic-scale analysis of bacterial gene and protein expression in the host. *Emerg Infect* **10**, 1357-62.

Boyce, J. D., Cullen, P. a, Nguyen, V., Wilkie, I. & Adler, B. (2006). Analysis of the *Pasteurella multocida* outer membrane sub-proteome and its response to the *in vivo* environment of the natural host. *Proteomics* **6**, 870-80.

Bracken, C. S., Baer, M. T., Abdur-Rashid, A., Helms, W. & Stojiljkovic, I. (1999). Use of heme-protein complexes by the *Yersinia enterocolitica* HemR receptor: histidine residues are essential for receptor function. *J Bacteriol* **181**, 6063-72.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Chem* **254**, 248-254.

Brennan, R. E., Corstvet, R. E. & Mcbride, J. W. (1997). Methods A simplified method for isolating outer membrane proteins from *Pasteurella haemolytica* A1. *J Microbiol Meth* **29**, 201-206.

Brickell, S. K., Thomas, L. M., Long, K. a, Panaccio, M. & Widders, P. R. (1998). Development of a PCR test based on a gene region associated with the pathogenicity of *Pasteurella multocida* serotype B:2, the causal agent of haemorrhagic septicaemia in Asia. *Vet Microbiol* **59**, 295-307.

Bridges, D. J., Pitt, A. R., Hanrahan, O., Brennan, K., Voorheis, H. P., Herzyk, P., de Koning, H. P. & Burchmore, R. J. S. (2008). Characterisation of the plasma membrane subproteome of bloodstream form *Trypanosoma brucei*. *Proteomics* **8**, 83-99.

Brockmeier, S. L. & Register, K. B. (2007). Expression of the dermonecrotic toxin by *Bordetella bronchiseptica* is not necessary for predisposing to infection with toxigenic *Pasteurella multocida*. *Vet Microbiol* **125**, 284-289.

Brown, T. A. (2007). *Genomes 3*, 3rd edn. US: Gerland Science Publishing.

Bulashevska, A. & Eils, R. (2006). Predicting protein subcellular locations using hierarchical ensemble of Bayesian classifiers based on Markov chains. *BMC Bioinformatics* **7**, 298.

Burchmore, R. (2006). Identification of anti-infective targets through comparative proteomics. *Expert Rev Anti Infect Ther* **4**, 163-5.

Buys, W. E. C. M., Smith, H. E., Kamps, A. M. I. E., Kamps, E. M. & Smits, M. A. (1990). Sequence of the dermonecrotic toxin of *Pasteurella multocida* ssp. *multocida*. *Nucleic Acids Res* **18**, 2815-16.

Capitini, C. M., Herrero, I. a, Patel, R., Ishitani, M. B. & Boyce, T. G. (2002). Wound infection with *Neisseria weaveri* and a novel subspecies of *Pasteurella multocida* in a child who sustained a tiger bite. *Clin Infect Dis* **34**, E74-6.

Carbone, A., Zinovyev, A. & Képès, F. (2005). Codon adaptation index as a measure of dominating codon bias. *Bioinformatics* **19**, 2005-2015.

Carmeli, C., Lifshitz, Y. & Friedberg, I. (1991). Spheroplast-derived membrane vesicles from *Rhodobacter capsulatus* cells catalyzing nucleotide transport. *Arch Biochem Biophys* **288**, 516-24.

Carpenter, T., Khalid, S. & Sansom, M. S. P. (2007). A multidomain outer membrane protein from *Pasteurella multocida*: modelling and simulation studies of PmOmpA. *Biochim Biophys Acta* **1768**, 2831-40.

Cescau, S., Cwerman, H., Létoffé, S., Delepelaire, P., Wandersman, C. & Biville, F. (2007). Heme acquisition by hemophores. *Biometals* **20**, 603-13.

Chanana, V., Majumdar, S., Ray, P., Sharma, M. & Rishi, P. (2006). Coordinated expression and immunogenicity of an outer membrane protein from *Salmonella enterica* serovar Typhi under iron limitation, oxidative stress and anaerobic conditions. *J Biomed Sci* **13**, 303-12.

Choi, C. & Chae, C. (2001). Enhanced detection of toxigenic *Pasteurella multocida* directly from nasal swabs using a nested polymerase chain reaction. *Vet J* **162**, 255-8.

Chooneea, D., Karlsson, R., Encheva, V., Arnold, C., Appleton, H. & Shah, H. (2010). Elucidation of the outer membrane proteome of *Salmonella enterica* serovar Typhimurium utilising a lipid-based protein immobilization technique. *BMC Microbiol* **10**, 44.

Christensen, H. & Bisgaard, M. (2006). The Genus *Pasteurella*. *Prokaryotes* **6**, 1062-1090.

Coenye, T. (2010). Response of sessile cells to stress: from changes in gene expression to phenotypic adaptation. *FEMS Immunol Med Mic* **59**, 239-52.

Coenye, T. & Nelis, H. J. (2010). In vitro and in vivo model systems to study microbial biofilm formation. *J Microbiol Meth* **83**, 89-105. Elsevier B.V.

Chung, J. Y., Zhang, Y. & Adler, B. (1998). The capsule biosynthetic locus of *Pasteurella multocida* A:1. *FEMS Microbiol Lett* **166**, 289-96.

Chung, J. W., Ng-Thow-Hing, C., Budman, L. I., Gibbs, B. F., Nash, J. H. E., Jacques, M. & Coulton, J. W. (2007). Outer membrane proteome of *Actinobacillus pleuropneumoniae*: LC-MS/MS analyses validate in silico predictions. *Proteomics* **7**, 1854-65.

Clock, S. a, Planet, P. J., Perez, B. a & Figurski, D. H. (2008). Outer membrane components of the Tad (tight adherence) secreton of *Aggregatibacter actinomycetemcomitans*. *J Bacteriol* **190**, 980-90.

Cole, K., Donoghue, a M., Reyes-Herrera, I., Rath, N. & Donoghue, D. J. (2006). Efficacy of iron chelators on *Campylobacter* concentrations in turkey semen. *Poultry Sci* **85**, 1462-5.

Collins, R. F. & Derrick, J. P. (2007). Wza : a new structural paradigm for outer membrane secretory proteins ? *Trends Microbiol* **15**, 96-100.

Combet, C., Jambon, M., Deléage, G. & Geourjon, C. (2002). Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics* **18**, 213-4.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M. & Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674-6.

Coquet, L., Cosette, P., Dé, E., Galas, L., Vaudry, H., Rihouey, C., Lerouge, P., Junter, G.-A. & Jouenne, T. (2005). Immobilization induces alterations in the outer membrane protein pattern of *Yersinia ruckeri*. *J Proteome Res* **4**, 1988-98.

Cordwell, S. J. (2006). Technologies for bacterial surface proteomics. *Curr Opin Microbiol* **9**, 320-9.

Cordwell, S. J., Len, A. C. L., Touma, R. G., Scott, N. E., Falconer, L., Jones, D., Connolly, A., Crossett, B. & Djordjevic, S. P. (2008). Identification of membrane-associated proteins from Campylobacter jejuni strains using complementary proteomics technologies. *Proteomics* **8**, 122-39.

Costerton, J. W., Ingram, J. M. & Cheng, K. J. (1974). Structure and function of the cell envelope of Gram-negative bacteria. *Bacteriol Rev* **38**, 87-110.

Cowan, M. K. (2012). *Microbiology: a systems approach*, 3rd edn. US: McGraw Hill.

Cox, A. J., Hunt, M. L., Boyce, J. D. & Adler, B. (2003). Functional characterization of HgbB, a new hemoglobin binding protein of *Pasteurella multocida*. *Microb Pathogenesis* **34**, 287-296.

Cox, C. D. (1986). Relationship between oxygen and siderophore synthesis in *Pseudomonas aeruginosa*. *Curr Microbiol* **14**, 19-23.

Cullen, P. A., Cordwell, S. J., Bulach, D. M., Haake, D. A. & Adler, B. (2002). Global analysis of outer membrane proteins from *Leptospira interrogans* serovar Lai. *Infect Immun* **70**, 2311-2318.

Cuthbertson, L., Mainprize, I. L., Naismith, J. H. & Whitfield, C. (2009). Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiol Mol Biol Rev* **73**, 155-77.

Van Cutsem, E., Simonart, G., Degand, H., Faber, A.-M., Morsomme, P. & Boutry, M. (2011). Gel-based and gel-free proteomic analysis of *Nicotiana tabacum* trichomes identifies proteins involved in secondary metabolism and in the (a)biotic stress response. *Proteomics* **11**, 440-54.

Dabo, S., Confer, A. W. & Quijano-Blas, R. A. (2003). Molecular and immunological characterization of *Pasteurella multocida* serotype A:3 OmpA: evidence of its role in *P. multocida* interaction with extracellular matrix molecules. *Microb Pathogenesis* **35**, 147-157.

Dabo, S. M., Taylor, J. D. & Confer, a W. (2007). *Pasteurella multocida* and bovine respiratory disease. *Anim Health Res Rev* **8**, 129-50.

Darveau, R. P., Charnetzky, W. T. & Hurlbert, R. E. (1980). Outer membrane protein composition of Yersinia pestis at different growth stages and incubation temperatures. *J Bacteriol* **143**, 942-9.

Davie, J. J. & Campagnari, A. a. (2009). Comparative proteomic analysis of the *Haemophilus ducreyi* porin-deficient mutant 35000HP::P2AB. *J Bacteriol* **191**, 2144-52.

Davies, R. L., Parton, R., Coote, J. G., Gibbs, H. a & Freer, J. H. (1992). Outer-membrane protein and lipopolysaccharide variation in *Pasteurella haemolytica* serotype A1 under different growth conditions. *J Gen Microbiol* **138**, 909-22.

Davies, R. L., Watson, J. & Caffrey, B. (2003a). Comparative analyses of *Pasteurella multocida* strains associated with the ovine respiratory and vaginal tracts. *Vet Rec* **152**, 7-10.

Davies, R. L. & Lee, I. (2004). Sequence diversity and molecular evolution of the heat-modifiable outer membrane protein gene (*ompA*) of *Mannheimia* (*Pasteurella*) *haemolytica*, *Mannheimia glucosida*, and *Pasteurella trehalosi*. *J Bacteriol* **186**, 5741-5752.

Davies, R. L., Whittam, T. S. & Selander, R. K. (2001). Sequence diversity and molecular evolution of the leukotoxin (*lktA*) gene in bovine and ovine strains of *Mannheimia* (*Pasteurella*) *haemolytica*. *J Bacteriol* **183**, 1394-1404.

Davies, R. L., MacCorquodale, R. & Caffrey, B. (2003b). Diversity of avian *Pasteurella multocida* strains based on capsular PCR typing and variation of the OmpA and OmpH outer membrane proteins. *Vet Microbiol* **91**, 169-82.

Davies, R. L., MacCorquodale, R. & Reilly, S. (2004). Characterisation of bovine strains of *Pasteurella multocida* and comparison with isolates of avian, ovine and porcine origin. *Vet Microbiol* **99**, 145-58.

Davies, R. L., MacCorquodale, R., Baillie, S. & Caffrey, B. (2003c). Characterization and comparison of *Pasteurella multocida* strains associated with porcine pneumonia and atrophic rhinitis. *J Med Microbiol* **52**, 59-67.

DeAngelis, P. L., Gunay, N. S., Toida, T., Mao, W.-jun & Linhardt, R. J. (2002). Identification of the capsular polysaccharides of Type D and F *Pasteurella multocida* as unmodified heparin and chondroitin, respectively. *Carbohyd Res* **337**, 1547-52.

Dear, P. H. (2007). *Bioinformatics*. Scion Publishing Ltd.

Dehio, C., Gray-owen, S. D. & Meyer, T. F. (1998). The role of neisserial Opa proteins in interactions with host cells. *Trends Microbiol* **6**, 489-495.

Deslandes, V., Nash, J. H. E., Harel, J., Coulton, J. W. & Jacques, M. (2007). Transcriptional profiling of *Actinobacillus pleuropneumoniae* under iron-restricted conditions. *BMC Genomics* **8**, 72.

Dey, S., Singh, V. P., Kumar, A. A., Sharma, B., Srivastava, S. K. & Singh, N. (2007). Comparative sequence analysis of 16S rRNA gene of *Pasteurella multocida* serogroup B isolates from different animal species. *Res Vet Sci* **83**, 1-4.

Diallo, I. S. & Frost, a J. (2000). Survival of avian strains of *Pasteurella multocida* in chicken serum. *Vet Microbiol* **72**, 153-61.

Djordjevic, S. P., Eamens, G. J., Ha, H. & Chin, J. C. (1998). Demonstration that Australian *Pasteurella multocida* isolates from sporadic outbreaks of porcine pneumonia are non-toxigenic (*toxA*-) and display heterogeneous DNA restriction endonuclease profiles compared with toxigenic isolates from herds w. *J Med Microbiol* **47**, 679-688.

Donachie, W., Lainson, F. A. & Hodgson, J. C. (1995). *Haemophilus, Actinobacillus* and *Pasteurella*. US: Plenum Publishing Corporation

Donnio, P. Y., Lerestif-Gautier, A. L. & Avril, J. L. (2004). Characterization of *Pasteurella* spp. strains isolated from human infections. *J Comp Path* **130**, 137-142.

Dumetz, F., Duchaud, E., Claverol, S., Orieux, N., Papillon, S., Lapaillerie, D. & Henaff, M. L. (2008). Analysis of the *Flavobacterium psychrophilum* outer-membrane subproteome and identification of new antigenic targets for vaccine by immunomics. *Microbiol* **154**, 1793-1801.

Dziva, F., Christensen, H., van Leengoed, L. A. M. G., Mohan, K. & Olsen, J. E. (2004). Differentiation of *Pasteurella multocida* isolates from cases of atrophic rhinitis in pigs from Zimbabwe by RAPD and ribotyping. *Vet Microbiol* **102**, 117-22.

Dziva, F., Muhairwa, A. P., Bisgaard, M. & Christensen, H. (2008). Diagnostic and typing options for investigating diseases associated with *Pasteurella multocida*. *Vet Microbiol* **128**, 1-22.

Díaz-Mejía, J. J., Babu, M. & Emili, A. (2009). Computational and experimental approaches to chart the *Escherichia coli* cell-envelope-associated proteome and interactome. *FEMS Microbiol Rev* **33**, 66-97.

De Vriendt, K., Theunissen, S., Carpentier, W., De Smet, L., Devreese, B. & Van Beeumen, J. (2005). Proteomics of Shewanella oneidensis MR-1 biofilm reveals differentially expressed proteins, including AggA and RibB. *Proteomics* **5**, 1308-16.

Ebanks, R. O., Dacanay, A., Goguen, M., Pinto, D. M. & Ross, N. W. (2004). Differential proteomic analysis of *Aeromonas salmonicida* outer membrane proteins in response to low iron and *in vivo* growth conditions. *Proteomics* **4**, 1074-85.

Eijkelkamp, B. a, Hassan, K. a, Paulsen, I. T. & Brown, M. H. (2011). Investigation of the human pathogen *Acinetobacter baumannii* under iron limiting conditions. *BMC Genomics* **12**, 126. BioMed Central Ltd.

Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Molec Biol* **179**, 125-42.

Ekins, A., Khan, A. G., Shouldice, S. R. & Schryvers, A. B. (2004). Lactoferrin receptors in gram-negative bacteria: insights into the iron acquisition process. *Biometals* **17**, 235-43.

Ellis, T. N. & Kuehn, M. J. (2010). Virulence and immunomodulatory roles of bacterial outer membrane vesicles. *Microbiol Mol Biol Rev* **74**, 81-94.

Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* **2**, 953-71.

Engelman, D. M., Steitz, T. A. & Goldman, A. (1986). Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu Rev Biophys Bio* **15**, 321-53.

Eroshenko, G. a, Vidyaeva, N. a & Kutyrev, V. V. (2010). Comparative analysis of biofilm formation by main and nonmain subspecies *Yersinia pestis* strains. *FEMS Immunol Med Microbiol* **59**, 513-20.

Eswar, N., Webb, B., Marti-Renom, M. a, Madhusudhan, M. S., Eramian, D., Shen, M.-Y., Pieper, U. & Sali, A. (2006). Comparative protein structure modeling using MODELLER. In *Current protocols in bioinformatics*, pp. 5.6.1-5.6.30. John Wiley & Sons Ltd.

Eswarappa, S. M., Janice, J., Nagarajan, A. G., Balasundaram, S. V., Karnam, G., Dixit, N. M. & Chakravortty, D. (2008). Differentially evolved genes of *Salmonella* pathogenicity islands : insights into the mechanism of host specificity in *Salmonella*. *PloS one* **3**, e3829.

Evans, K. & Poole, K. (1999). The MexA-MexB-OprM multidrug efflux system of *Pseudomonas aeruginosa* is growth-phase regulated. *FEMS Microbiol Lett* **173**, 35-9.

Ewers, C., Lübke-Becker, A., Bethe, A., Kiebling, S., Filter, M. & Wieler, L. H. (2006). Virulence genotype of *Pasteurella multocida* strains isolated from different hosts with various disease status. *Vet Microbiol* **114**, 304-17.

Facey, S. J. & Kuhn, A. (2010). Biogenesis of bacterial inner-membrane proteins. *Cell Mol Life Sci* **67**, 2343-62.

Faraldo-Gómez, J. D. & Sansom, M. S. P. (2003). Acquisition of siderophores in gram-negative bacteria. *Nat Rev Mol Cell Bio* **4**, 105-16.

Fariselli, P., Finocchiaro, G. & Casadio, R. (2003). SPEPlip: the detection of signal peptide and lipoprotein cleavage sites. *Bioinformatics* **19**, 2498-2499.

Figueira, C. P., Huerre, M., Ristow, P., Bourhy, P., Ave, P., Girons, I. S., Ko, A. I. & Picardeau, M. (2007). The OmpA-like protein Loa22 is essential for Leptospiral virulence. *PLoS Pathog* **3**.

Filip, C., Fletcher, G., Wulff, J. L. & Earhart, C. F. (1973). Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium-lauryl sarcosinate. *J Bacteriol* **115**, 717-22.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., & other authors. **(2010).** The Pfam protein families database. *Nucleic Acids Res* **38**, D211-22.

Ford, R. C., Brunkan-LaMontagne, A. L., Collins, R. F., Clarke, B. R., Harris, R., Naismith, J. H. & Whitfield, C. (2009). Structure-function relationships of the outer membrane translocon Wza investigated by cryo-electron microscopy and mutagenesis. *J Struct Biol* **166**, 172-82. Elsevier Inc.

Frasch, C. E., McNelis, R. M. & Gotschlich, E. C. (1976). Strain-specific variation in the protein and lipopolysaccharide composition of the group B meningococcal outer membrane. *J Bacteriol* **127**, 973-81.

Freinkman, E., Chng, S. & Kahne, D. (2010). The complex that inserts lipopolysaccharide into the bacterial outer membrane forms. *PNAS*.

Freinkman, E., Chng, S.-S. & Kahne, D. (2011). The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel. *PNAS* **108**, 2486-91.

Futoma-Kołoch, B., Bugla-Płoskońska, G., Doroszkiewicz, W. & Kaca, W. (2006). Survival of *Proteus mirabilis* O3 (S1959), O9 and O18 strains in normal human serum (NHS) correlates with the diversity of their outer membrane proteins (OMPs). *Polish J Microbiol* **55**, 153-6.

Fyshe, A., Liu, Y., Szafron, D., Greiner, R. & Lu, P. (2008). Improving subcellular localization prediction using text classification and the gene ontology. *Bioinformatics* **24**, 2512-7.

Gaddy, J. a, Tomaras, A. P. & Actis, L. a. **(2009).** The *Acinetobacter baumannii* 19606 OmpA protein plays a role in biofilm formation on abiotic surfaces and in the interaction of this pathogen with eukaryotic cells. *Infect Immun* **77**, 3150-60.

Gallot-lavalme, T., Zygmunt, M. S. & Btzani, A. C. G. (1995). Growth phase-dependent variations in protein profile of *Brucella melitensis*. *Res Microbiol* 146, 227-36.

Gao, Q.-B., Ye, X.-F., Jin, Z.-C. & He, J. (2010). Improving discrimination of outer membrane proteins by fusing different forms of pseudo amino acid composition. *Anal Biochem* **398**, 52-9.

Gardy, J. L., Laird, M. R., Chen, F., Rey, S., Walsh, C. J., Ester, M. & Brinkman, F. S. L. (2005). PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* **21**, 617-23.

Gardy, J. L., Spencer, C., Wang, K., Ester, M., Tusnady, G. E., Simon, I., Hua, S., DeFays, K., Lambert, C., & other authors. **(2003).** PSORT-B: improving protein subcellular localization prediction for Gram-negative bacteria. *Nucleic Acids Res* **31**, 3613-3617.

Gardy, J. L. & Brinkman, F. S. L. (2006). Methods for predicting bacterial protein subcellular localization. *Nat Rev Microbiol* **4**, 741-751.

Garrido, M. E., Bosch, M., Medina, R., Bigas, A., Llagostera, M., Perez de Rozas, A. M., Badiola, I. & Barbe, J. (2003). *fur*-independent regulation of the *Pasteurella multocida hbpA* gene encoding a haemin-binding protein. *Microbiol* **149**, 2273-2281.

Garrow, A. G., Agnew, A. & Westhead, D. R. (2005a). TMB-Hunt: a web server to screen sequence sets for transmembrane beta-barrel proteins. *Nucleic Acids Res* **33**, W188-92.

Garrow, A. G., Agnew, A. & Westhead, D. R. (2005b). TMB-Hunt : An amino acid composition based method to screen proteomes for beta-barrel transmembrane proteins. *BMC Bioinformatics* **16**, 1-16.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). Protein identification and analysis tools in the ExPASy server. In *The proteomics protocols handbook*, pp. 531-52. Edited by J. M. Walker. Humana Press.

Gerlach, R. G. & Hensel, M. (2007). Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *Int J Med Microbiol* **297**, 401-15.

Gautam, R., Kumar, a a, Singh, V. P., Singh, V. P., Dutta, T. K. & Shivachandra, S. B. (2004). Specific identification of *Pasteurella multocida* serogroup-A isolates by PCR assay. *Res Vet Sci* **76**, 179-85.

Glorioso, J. C., Jones, G. W., Rush, H. G., Pentler, L. J., Darif, A. & Coward, J. O. E. E. (1982). Adhesion of type A *Pasteurella multocida* to rabbit pharyngeal cells and its possible role in rabbit respiratory tract infections. *Infect Immun* **35**, 1103-1109.

Goudenège, D., Avner, S., Lucchetti-Miganeh, C. & Barloy-Hubler, F. (2010). CoBaltDB: Complete bacterial and archaeal orfeomes subcellular localization database and associated resources. *BMC Microbiol* **10**, 88.

Grenier, D., Goulet, V. & Mayrand, D. (2001). The capacity of *Porphyromonas gingivalis* to multiply under iron-limiting conditions correlates with its pathogenicity in an animal model. *J Dent Res* **80**, 1678-1682.

Gromiha, M. M. (2005). Motifs in outer membrane protein sequences: applications for discrimination. *Biophys Chem* **117**, 65-71.

Gromiha, M. M. & Suwa, M. (2005). A simple statistical method for discriminating outer membrane proteins with better accuracy. *Bioinformatics* **21**, 961-8.

Gromiha, M. M. & Suwa, M. (2006a). Influence of amino acid properties for discriminating outer membrane proteins at better accuracy. *Biochim Biophys Acta* **1764**, 1493-7.

Gromiha, M. M. & Suwa, M. (2006b). Discrimination of outer membrane proteins using membrane learning algorithms. *Proteins* **63**, 1031-1037.

Gromiha, M. M. & Yabuki, Y. (2008). Functional discrimination of membrane proteins using machine learning techniques. *BMC Bioinformatics* **9**, 135.

Gromiha, M. M., Ahmad, S. & Suwa, M. (2005). TMBETA-NET: discrimination and prediction of membrane spanning beta-strands in outer membrane proteins. *Nucleic Acids Res* **33**, W164-7.

Gromiha, M. M., Yabuki, Y. & Suwa, M. (2007). TMB finding pipeline: novel approach for detecting beta-barrel membrane proteins in genomic sequences. *J Chem Inf Model* **47**, 2456-61.

Gromiha, M. M., Yabuki, Y., Suresh, M. X., Thangakani, A. M., Suwa, M. & Fukui, K. (2009). TMFunction : database for functional residues in membrane proteins. *Nucleic Acids Res* **37**, 201-204.

Gunawardana, G. A., Townsend, K. M. & Frost, A. J. (2000). Molecular characterisation of avian *Pasteurella multocida* isolates from Australia and Vietnam by REP-PCR and PFGE. *Vet Microbiol* **72**, 97-109.

Gupta, N. & Pevzner, P. A. (2009). False discovery rates of protein identifications: a strike against the two-peptide rule. *J Proteome Res* **8**, 4173-4181.

van Haagen, H. H. H. B. M., 't Hoen, P. a C., de Morrée, A., van Roon-Mom, W. M. C., Peters, D. J. M., Roos, M., Mons, B., van Ommen, G.-J. & Schuemie, M. J. (2011). In silico discovery and experimental validation of new protein-protein interactions. *Proteomics* **11**, 843-53.

Hancock, I. & Poxton, I. (1988). *Bacterial cell surface techniques*. Wiley-VCH.

Hari-Dass, R., Shah, C., Meyer, D. J. & Raynes, J. G. (2005). Serum amyloid A protein binds to outer membrane protein A of gram-negative bacteria. *J Biol Chem* **280**, 18562-7.

Harper, M., Boyce, J. D. & Adler, B. (2006). *Pasteurella multocida* pathogenesis : 125 years after Pasteur. *FEMS Microbiol Lett* **265**, 1-10.

Harper, M., Boyce, J. D., Cox, A. D., St Michael, F., Wilkie, I. W., Blackall, P. J. & Adler, B. (2007a). *Pasteurella multocida* expresses two lipopolysaccharide glycoforms simultaneously, but only a single form is required for virulence: identification of two acceptor-specific heptosyl I transferases. *Infect Immun* **75**, 3885-93.

Harper, M., Cox, A. D., St Michael, F., Parnas, H., Wilkie, I., Blackall, P. J., Adler, B. & Boyce, J. D. (2007b). Decoration of *Pasteurella multocida* lipopolysaccharide with phosphocholine is important for virulence. *J Bacteriol* **189**, 7384-91.

Hatfaludi, T., Al-Hasani, K., Dunstone, M., Boyce, J. & Adler, B. (2008). Characterization of TolC efflux pump proteins from *Pasteurella multocida*. *Antimicrob Agents Chemother* **52**, 4166-71.

Hatfaludi, T., Al-Hasani, K., Boyce, J. D. & Adler, B. (2010). Outer membrane proteins of *Pasteurella multocida*. *Vet Microbiol* **144**, 1-17.

Hauberg, L., Schmidt, F., Scharf, C., Dörr, J., Völker, U. & Reinhold-Hurek, B. (2010). Proteomic characterization of a *pilR* regulatory mutant of *Azoarcus* sp. strain BH72 with the aid of gel-based and gel-free approaches. *Proteomics* **10**, 458-69.

Hays, J. P., Selm, S. V., Hoogenboezem, T., Esteva, S., Eadie, K., Veelen, P. V., Tommassen, J., Belkum, A. V. & Hermans, P. W. M. (2005). Identification and characterization of a novel outer membrane protein (OmpJ) of *Moraxella catarrhalis* that exists in two major forms. *J Bacteriol* **187**, 7977-7984.

Heim, S., Ferrer, M., Heuer, H., Regenhardt, D., Nimtz, M. & Timmis, K. N. (2003). Proteome reference map of *Pseudomonas putida* strain KT2440 for genome expression profiling : distinct responses of KT2440 and *Pseudomonas aeruginosa* strain PAO1 to iron deprivation and a new form of superoxide dismutase. *Environ Microbiol* **5**, 1257-1269.

Heinz, E., Tischler, P., Rattei, T., Myers, G., Wagner, M. & Horn, M. (2009). Comprehensive *in silico* prediction and analysis of chlamydial outer membrane proteins reflects evolution and life style of the *Chlamydiae*. *BMC Genomics* **10**, 1-18.

Heli, H., Mirtorabi, S. & Karimian, K. (2011). Advances in iron chelation : an update. *Expert Opin Ther Pat* **21**, 819-856.

Henningsen, R., Gale, B. L., Straub, K. M. & DeNagel, D. C. (2002). Application of zwitterionic detergents to the solubilization of integral membrane proteins for two-dimensional gel electrophoresis and mass spectrometry. *Proteomics* **2**, 1479-88.

Henry, I. & Sharp, P. M. (2007). Predicting gene expression level from codon usage bias. *Mol Biol Evol* **24**, 10-2.

Hildebrand, D., Heeg, K. & Kubatzky, K. F. (2011). *Pasteurella multocida* toxin-stimulated osteoclast differentiation is B cell dependent. *Infect Immun* **79**, 220-8.

Hirsh, D. C., Jessup, D. A., Snipes, K. P., Carpenter, T. E., Hird, W. & Mccapes, H. (1990). Characteristics of *Pasteurella multocida* isolated from waterfowl and associated avian species in California. *J Wildlife Dis* **26**, 204-209.

Ho, W., Yu, R. C. & Chou, C. C. (2004). Effect of iron limitation on the growth and cytotoxin production of *Salmonella choleraesuis* SC-5. *Int J Food Microbiol* **90**, 295-302.

Hobb, R. I., Fields, J. a, Burns, C. M. & Thompson, S. a. **(2009).** Evaluation of procedures for outer membrane isolation from *Campylobacter jejuni*. *Microbiol* **155**, 979-88.

Holmes, K., Mulholland, F., Pearson, B. M., Pin, C., McNicholl-Kennedy, J., Ketley, J. M. & Wells, J. M. (2006). *Campylobacter jejuni* gene expression in response to iron limitation and the role of Fur. *Microbiol* **151**, 243-57.

Hong, H., Patel, D. R., Tamm, L. K. & van den Berg, B. (2006). The outer membrane protein OmpW forms an eight-stranded beta-barrel with a hydrophobic channel. *J Biol Chem* **281**, 7568-77.

Horler, R. S. P., Butcher, A., Papangelopoulos, N., Ashton, P. D. & Thomas, G. H. (2009). EchoLOCATION : an in silico analysis of the subcellular locations of *Escherichia coli* proteins and comparison with experimentally derived locations. *Bioinformatics* **25**, 163-166.

Hu, J. & Yan, C. (2008). A method for discovering transmembrane beta-barrel proteins in Gram-negative bacterial proteomes. *Comput Biol Chem* **32**, 298-301.

Hu, Q., Han, X., Zhou, X., Ding, S., Ding, C. & Yu, S. (2010). Characterization of biofilm formation by *Riemerella anatipestifer*. *Vet Microbiol* **144**, 429-36. Elsevier B.V.

Hua, S. & Sun, Z. (2001). Support vector machine approach for protein subcellular localization prediction. *Bioinformatics* **17**, 721-8.

Hunt, M. L., Adler, B. & Townsend, K. M. (2000). The molecular biology of *Pasteurella multocida*. *Vet Microbiol* **72**, 3-25.

Huntley, J. F., Conley, P. G., Hagman, K. E. & Norgard, M. V. (2007). Characterization of *Francisella tularensis* outer membrane proteins. *J Bacteriol* **189**, 561-74.

Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. *Mol Biol Evol* **2**, 13-34.

Imai, K., Asakawa, N., Tsuji, T., Akazawa, F., Ino, A. & Sonoyama, M. (2008). Bioinformation SOSUI-GramN : high performance prediction for sub-cellular localization of proteins in Gram- negative bacteria. *Bioinformation* **2**, 417-421.

Isaacson, R. E. & Trigo, E. (1995). Pili of *Pasteurella multocida* of porcine origin. *FEMS Microbiol Lett* **132**, 247-51.

Jabbari, A. R., Esmaelzadeh, M. & Moazeni Jula, G. R. (2006). Polymerase chain reaction typing of *Pasteurella multocida* capsules isolated in Iran. *Iran J Vet Res* **7**, 50-55.

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.

Jackups, R., Cheng, S. & Liang, J. (2006). Sequence motifs and antimotifs in beta-barrel membrane proteins from a genome-wide analysis: the Ala-Tyr dichotomy and chaperone binding motifs. *J Mol Biol* **363**, 611-23.

Jacob-Dubuisson, F., Locht, C. & Antoine, R. (2001). Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. *Mol Microbiol* **40**, 306-13.

Jacques, M., Kobisch, M., Bélanger, M. & Dugal, F. (1993). Virulence of capsulated and noncapsulated isolates of *Pasteurella multocida* and their adherence to porcine respiratory tract cells and mucus. *Infect Immun* **61**, 4785-92.

Jacques, M., Bélanger, M., Diarra, M. S., Dargis, M. & Malouin, F. (1994). Modulation of *Pasteurella multocida* capsular polysaccharide during growth under iron-restricted conditions and *in vivo*. *Microbiol* **140 (**, 263-70.

Jacques, M., Aragon, V. & Tremblay, Y. D. N. (2010). Biofilm formation in bacterial pathogens of veterinary importance. *Anim Health Res Rev* **11**, 97-121.

Jaglic, Z., Kucerova, Z., Nedbalcova, K., Pavlik, I., Alexa, P. & Bartos, M. (2005). Characterisation and comparison of *Pasteurella multocida* isolated from different species in the Czech Republic : capsular PCR typing , ribotyping and dermonecrotoxin production. *Vet Med* **8**, 345-354.

Jaglic, Z., Kucerova, Z., Nedbalcova, K., Kulich, P. & Alexa, P. (2006). Characterisation of *Pasteurella multocida* isolated from rabbits in the Czech Republic. *Vet Med-Czech* **5**, 278-287.

Jain, A. & Agarwal, A. (2009). Biofilm production, a marker of pathogenic potential of colonizing and commensal staphylococci. *J Microbiol Meth* **76**, 88-92.

Janin, J. (1979). Surface and inside volumes in globular proteins. *Nature* **277**, 491-492.

Jarosławski, S., Duquesne, K., Sturgis, J. N. & Scheuring, S. (2009). High-resolution architecture of the outer membrane of the Gram-negative bacteria *Roseobacter denitrificans. Mol Microbiol* **74**, 1211-22.

Juncker, A. S., Willenbrock, H., Heijne, G. V. O. N., Brunak, S., Nielsen, H. & Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **12**, 1652-1662.

Jungblut, P. R., Schiele, F., Zimny-Arndt, U., Ackermann, R., Schmid, M., Lange, S., Stein, R. & Pleissner, K.-P. (2010). *Helicobacter pylori* proteomics by 2-DE/MS, 1-DE-LC/MS and functional data mining. *Proteomics* **10**, 182-93.

Karunakaran, S., Nair, G. K. & Mini, M. (2009). REP-PCR profiles of *Pasteurella multocida* isolates from cases of haemorrhagic septicaemia. *Tamilnadu J Vet Anim Sci* **5**, 64-66.

Katta, A. V. S. K. M., Marikkannu, R., Basaiawmoit, R. V. & Krishnaswamy, S. (2004). Consensus based validation of membrane porins. *In silico Biol* **4**.

Kaur, R., Sharma, A., Majumdar, S., Ganguly, N. K. & Chakraborti, A. (2003). Outer-membrane-protein subtypes of *Haemophilus influenzae* isolates from North India. *J Med Microbiol* **52**, 693-696.

Keedwell, E. & Narayanan, A. (2005). *Intelligent bioinformatics:The application of artificial intelligence techniques to bioinformatics problems.* UK: John Wiley & Sons Ltd.

Khemiri, A., Galland, A., Vaudry, D., Chan Tchi Song, P., Vaudry, H., Jouenne, T. & Cosette, P. (2008). Outer-membrane proteomic maps and surface-exposed proteins of *Legionella pneumophila* using cellular fractionation and fluorescent labelling. *Anal Bioanal Chem* **390**, 1861-71.

Kiefer, F., Arnold, K., Künzli, M., Bordoli, L. & Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* **37**, D387-92.

Kim, S., Kim, H., Reuhs, B. L. & Mauer, L. J. (2006). Differentiation of outer membrane proteins from *Salmonella enterica* serotypes using Fourier transform infrared spectroscopy and chemometrics. *Lett Appl Microbiol* **42**, 229-34.

Kimizuka, R., Kato, T., Hashimoto, S., Yamanaka-Okada, A., Okuda, K. & Ishihara, K. (2009). Congo red-binding protein in rough-phenotype *Aggregatibacter actinomycetemcomitans* is amyloid-like fiber. *Bull Tokyo Dent Coll* **50**, 23-9.

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. BioMed Central Ltd.

Knobloch, J. K.-M., Horstkotte, M. a, Rohde, H. & Mack, D. (2002). Evaluation of different detection methods of biofilm formation in *Staphylococcus aureus*. *Med Microbiol Immun* **191**, 101-6.

Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* **7**, 206-14.

Kolberg, J., Hammerschmidt, S., Frank, R., Jonák, J., Sanderová, H. & Aase, A. (2008). The surface-associated elongation factor Tu is concealed for antibody binding on viable pneumococci and meningococci. *FEMS Immunol Med Mic* **53**, 222-30.

Komatsuzawa, H., Asakawa, R., Kawai, T., Ochiai, K., Fujiwara, T., Taubman, M. a, Ohara, M., Kurihara, H. & Sugai, M. (2002). Identification of six major outer membrane proteins from *Actinobacillus actinomycetemcomitans*. *Gene* **288**, 195-201.

Kouyianou, K., Aivaliotis, M., Gavaert, K., Karas, M. & Tsiotis, G. (2010). Membrane proteome of the green sulfur bacterium *Chlorobium tepidum* (syn . *Chlorobaculum tepidum*) analyzed by gel-based and gel-free methods. *Photosynth Res* **104**, 153-162.

Krewulak, K. D. & Vogel, H. J. (2008). Structural biology of bacterial iron uptake. *Biochim Biophys Acta* **1778**, 1781-804.

Krogh, a, Larsson, B., von Heijne, G. & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567-80.

Kuehn, M. J. & Kesty, N. C. (2005). Bacterial outer membrane vesicles and the host-pathogen interaction. *Genes Dev* **19**, 2645-2655.

Kuhnert, P. & Christensen, H. (2008). Pasteurellaceae *biology, genomics and molecular aspects.* Norfolk, UK: Caister Academic Press.

Kuhnert, P. & Korczak, B. M. (2006). Prediction of whole-genome DNA-DNA similarity, determination of G+C content and phylogenetic analysis within the family *Pasteurellaceae* by multilocus sequence analysis (MLSA). *Microbiol* **152**, 2537-48.

Kumar, A. A., Shivachandra, S. B., Biswas, A., Singh, V. P. & Singh, V. P. (2004). Prevalent serotypes of *Pasteurella multocida* isolated from different animal and avian species in India. *Vet Res* **28**, 657-667.

Kumar, G., Sharma, P., Rathore, G., Bisht, D. & Sengupta, U. (2009a). Proteomic analysis of outer membrane proteins of *Edwardsiella tarda*. *J Appl Microbiol* **108**, 2214-2221.

Kumar, P., Singh, V. P., Agrawal, R. K. & Singh, S. (2009b). Identification of *Pasteurella multocida* isolates of ruminant origin using polymerase chain reaction and their antibiogram study. *Trop Anim Health Prod* **41**, 573-578.

Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-32.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Landini, P. (2009). Cross-talk mechanisms in biofilm formation and responses to environmental and physiological stress in *Escherichia coli*. *Res Microbiol* **160**, 259-66.

Lawrence, P. K., Kittichotirat, W., McDermott, J. E. & Bumgarner, R. E. (2010). A three-way comparative genomic analysis of *Mannheimia haemolytica* isolates. *BMC genomics* **11**, 535.

Lee, E.-Y., Bang, J. Y., Park, G. W., Choi, D.-S., Kang, J. S., Kim, H.-J., Park, K.-S., Lee, J.-O., Kim, Y.-K., & other authors. **(2007).** Global proteomic profiling of native outer membrane vesicles derived from *Escherichia coli*. *Proteomics* **7**, 3143-53.

Lee, E.-young, Choi, D.-sic, Kim, K.-pyo, Gho, Y. S. & Introduction, I. (2008). Proteomics in Gram-negative bacterial outer membrane vesicles. *Mass Spectrom Rev* **27**, 535- 555.

Lenco, J., Hubálek, M., Larsson, P., Fucíková, A., Brychta, M., Macela, A. & Stulík, J. (2007). Proteomics analysis of the *Francisella tularensis* LVS response to iron restriction: induction of the *F. tularensis* pathogenicity island proteins IglABC. *FEMS Microbiol Lett* **269**, 11-21.

Letunic, I., Doerks, T. & Bork, P. (2009). SMART 6: recent updates and new developments. *Nucleic Acids Res* **37**, D229-32.

Leys, D., Backers, K., Meyer, T. E., Hagen, W. R., Cusanovich, M. A. & Beeumen, J. J. V. (2000). Crystal structures of an oxygen-binding cytochrome c from *Rhodobacter sphaeroides*. *J Biol Chem* **275**, 16050 -16056.

Lichtensteiger, C. a, Steenbergen, S. M., Lee, R. M., Polson, D. D. & Vimr, E. R. (1996). Direct PCR analysis for toxigenic *Pasteurella multocida*. *J Clin Microbiol* **34**, 3035-9.

Lin, J., Huang, S. & Zhang, Q. (2002). Outer membrane proteins: key players for bacterial adaptation in host niches. *Microbes Infect* **4**, 325-31.

Lin, X.-min, Wu, L.-na, Li, H., Wang, S.-ying & Peng, X.-xian. **(2008).** Downregulation of Tsx and OmpW and upregulation of OmpX are required for iron homeostasis in *Escherichia coli*. *J Proteome Res* **7**, 1235-43.

Linke, D., Riess, T., Autenrieth, I. B., Lupas, A. & Kempf, V. a J. (2006). Trimeric autotransporter adhesins: variable structure, common function. *Trends Microbiol* **14**, 264-70.

Liu, G. Y., Nie, P., Zhang, J. & Li, N. (2008). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of *Flavobacterium columnare*. *J Fish Dis* **31**, 269-276.

Lo, M., Murray, G. L., Khoo, C. A., Haake, D. a, Zuerner, R. L. & Adler, B. (2010). Transcriptional response of *Leptospira interrogans* to iron limitation and characterization of a PerR homolog. *Infect Immun* **78**, 4850-9.

Lomize, A. L., Pogozheva, I. D., Lomize, M. A. & Mosberg, H. I. (2006). Positioning of proteins in membranes : A computational approach Positioning of proteins in membranes : A computational approach. *Protein Sci* **15**, 1318-1333.

Lorian, V. (1989). In vitro simulation of in vivo conditions: physical state of the culture medium. *J Clin Microbiol* **27**, 2403-6.

Lu, Z., Szafron, D., Greiner, R., Lu, P., Wishart, D. S., Poulin, B., Anvik, J., Macdonell, C. & Eisner, R. (2004). Predicting subcellular localization of proteins using machine-learned classifiers. *Bioinformatics* **20**, 547-556.

Lugtenberg, B., van Boxtel, R. & de Jong, M. (1984). Atrophic rhinitis in swine: correlation of *Pasteurella multocida* pathogenicity with membrane protein and lipopolysaccharide patterns. *Infect Immun* **46**, 48-54.

Luo, Y., Zeng, Q., Glisson, J. R., Jackwood, M. W., Cheng, I. H. & Wang, C. (1999). Sequence analysis of *Pasteurella multocida* major outer membrane protein (OmpH) and application of synthetic peptides in vaccination of chickens against homologous strain challenge. *Vaccine* **17**, 821-31.

Luscombe, N. M., Greenbaum, D. & Gerstein, M. (2001). What is bioinformatics? A proposed definition and overview of the field. *Methods Inform Med* **4**, 346-58.

Ma, Q. & Wood, T. K. (2009). OmpA influences *Escherichia coli* biofilm formation by repressing cellulose production through the CpxRA two-component system. *Environ Microbiol* **11**, 2735-46.

Marandi, M. & Mittal, K. R. (1996). Characterization of an outer membrane protein of *Pasteurella multocida* belonging to the OmpA family. *Vet Microbiol* **53**, 303-14.

March, C., Moranta, D., Regueiro, V., Llobet, E., Tomás, A., Garmendia, J. & Bengoechea, J. a. (2011). *Klebsiella pneumoniae* outer membrane protein A is required to prevent the activation of airway epithelial cells. *J Biol Chem* **286**, 9956-67.

Mariana, N. S., Salman, S. A., Neela, V. & Zamberi, S. (2009). Evaluation of modified Congo red agar for detection of biofilm produced by clinical isolates of methicillin – resistance *Staphylococcus aureus*. *J Microbiol* **3**, 330-338.

Markwell, M. A. K., Haas, S. M., Bieber, L. L. & Tolbert, N. E. (1978). Determination of the Lowry procedure to simplify protein in membrane and lipoprotein samples. *Anal Biochem* **87**, 206-210.

Maruvada, R. & Kim, K. S. (2011). Extracellular loops of the *Eschericia coli* outer membrane protein A contribute to the pathogenesis of meningitis. *JI Infect Dis* **203**, 131-40.

Mashburn-Warren, L., McLean, R. J. C. & Whiteley, M. (2008). Gram-negative outer membrane vesicles: beyond the cell surface. *Geobiology* **6**, 214-9.

Massé, E. & Arguin, M. (2005). Ironing out the problem: new mechanisms of iron homeostasis. *Trends Biochem Sci* **30**, 462-8.

Masuda, T., Tomita, M. & Ishihama, Y. (2008). Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J Proteome Res* **7**, 731-40.

May, B. J., Zhang, Q., Li, L. L., Paustian, M. L., Whittam, T. S. & Kapur, V. (2001). Complete genomic sequence of *Pasteurella multocida*, Pm70. *PNAS* **98**, 3460-5.

Mbuthia, P. G., Christensen, H., Boye, M., Majken, K., Petersen, D., Bisgaard, M., Nyaga, P. N. & Olsen, J. E. (2001). Specific detection of *Pasteurella multocida* in chickens with fowl cholera and in pig lung tissues using fluorescent rRNA in situ hybridization. *J Clin Microbiol* **39**, 2627-2633.

Melnikow, E., Dornan, S., Sargent, C., Duszenko, M., Evans, G., Gunkel, N., Selzer, P. M. & Ullrich, H. J. (2005). Microarray analysis of *Haemophilus parasuis* gene expression under in vitro growth conditions mimicking the in vivo environment. *Vet Microbiol* **110**, 255-63.

Melnikow, E., Schoenfeld, C., Spehr, V., Warrass, R., Gunkel, N., Duszenko, M., Selzer, P. M. & Ullrich, H. J. (2008). A compendium of antibiotic-induced transcription profiles reveals broad regulation of *Pasteurella multocida* virulence genes. *Vet Microbiol* **131**, 277-92.

Merlin, C., Gardiner, G., Durand, S. & Masters, M. (2002). The *Escherichia coli metD* locus encodes an ABC transporter which includes Abc (MetN), YaeE (MetI), and YaeC (MetQ). *J Bacteriol* **184**, 5513-5517.

Metruccio, M. M. E., Fantappiè, L., Serruto, D., Muzzi, A., Roncarati, D., Donati, C., Scarlato, V. & Delany, I. (2009). The Hfq-dependent small noncoding RNA NrrF directly mediates Fur-dependent positive regulation of succinate dehydrogenase in *Neisseria meningitidis*. *J Bacteriol* **191**, 1330-42.

Miethke, M. & Marahiel, M. a. **(2007).** Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol R* **71**, 413-51.

Miflin, J. K. & Blackall, P. J. (2001). Development of a 23S rRNA-based PCR assay for the identification of *Pasteurella multocida*. *Lett Appl Microbiol* **33**, 216-221.

Mikkelsen, H., Duck, Z., Lilley, K. S. & Welch, M. (2007). Interrelationships between colonies, biofilms, and planktonic cells of *Pseudomonas aeruginosa*. *J Bacteriol* **189**, 2411-6.

Miller, I., Wait, R., Sipos, W. & Gemeiner, M. (2009). A proteomic reference map for pig serum proteins as a prerequisite for diagnostic applications. *Res Vet Sci* **86**, 362-7.

Miot, M. & Betton, J.-M. (2004). Protein quality control in the bacterial periplasm. *Microb Cell Fact* **3**.

Mira, A., Martín-cuadrado, A. B., Auria, G. D. & Rodríguez-valera, F. (2010). The bacterial pan-genome : a new paradigm in microbiology. *Int Microbiol* **13**, 45-57.

Mirus, O. & Schleiff, E. (2005). Prediction of beta-barrel membrane proteins by searching for restricted domains. *BMC bioinformatics* **6**, 254.

Mizan, S., Henk, A., Stallings, A., Maier, M. & Lee, M. D. (2000). Cloning and characterization of sialidases with 2-6' and 2-3' sialyl lactose specificity from *Pasteurella multocida*. *J Bcteriol* **182**, 6874-83.

Mohan, K., Kelly, P. J., Hill, F. W. G., Muvavarirwa, P. & Pawandiwa, A. (1997). Phenotype and serotype of *Pasteurella multocida* isolates from diseases of dogs and cats in Zimbabwe. *Comp Immum Microbiol Infect Dis* **20**, 29-34.

Molloy, M. P., Herbert, B. R., Slade, M. B., Rabilloud, T., Nouwens, a S., Williams, K. L. & Gooley, a a. **(2000).** Proteomic analysis of the *Escherichia coli* outer membrane. *Eur J Biochem* **267**, 2871-81.

Mondal, U. K., Sur, S., Bothra, a K. & Sen, A. (2008). Comparative analysis of codon usage patterns and identification of predicted highly expressed genes in five *Salmonella* genomes. *Indian J Med Microbiol* **26**, 313-21.

Morona, R., Krämer, C. & Henning, U. (1985). Bacteriophage receptor area of outer membrane protein OmpA of *Escherichia coli* K-12. *J Bacteriol* **164**, 539-43.

Muhairwa, A. P., Christensen, J. P. & Bisgaard, M. (2002). Serum resistance of *Pasteurella multocida* in avian and porcine sera, and comparative virulence

investigations of selected serum-sensitive and resistant strains in chickens. *Avian Pathol* **31**, 183-91.

Mullen, L. M., Bossé, J. T., Nair, S. P., Ward, J. M., Rycroft, A. N., Robertson, G., Langford, P. R. & Henderson, B. (2008a). *Pasteurellaceae* ComE1 proteins combine the properties of fibronectin adhesins and DNA binding competence proteins. *PloS one* **3**, e3991.

Mullen, L. M., Nair, S. P., Ward, J. M., Rycroft, A. N., Williams, R. J., Robertson, G., Mordan, N. J. & Henderson, B. (2008b). Novel adhesin from *Pasteurella multocida* that binds to the integrin-binding fibronectin FnIII9-10 repeats. *Infect Immun* **76**, 1093-104.

Murphy, T. F. & Loeb, M. R. (1989). Isolation of the outer membrane of *Branhamella catarrhalis*. *Microb Pathogenesis* **6**, 159-174.

Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**, 536-40.

Najimi, M., Lemos, M. L. & Osorio, C. R. (2008). Identification of siderophore biosynthesis genes essential for growth of *Aeromonas salmonicida* under iron limitation conditions. *Appl Environ Microbiol* **74**, 2341-8.

Narita, S.-I., Matsuyama, S.-I. & Tokuda, H. (2004). Lipoprotein trafficking in *Escherichia coli*. *Archiv Microbiol* **182**, 1-6.

Natale, P., Brüser, T. & Driessen, A. J. M. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta* **1778**, 1735-56.

Negrete-Abascal, E., Tenorio, V. R. & de la Garza, M. (1999). Secretion of proteases from *Pasteurella multocida* isolates. *Curr Microbiol* **38**, 64-7.

Niven, D. F., Donga, J. & Archibald, F. S. (1989). Responses of *Haemophilus pleuropneumoniae* to iron restriction: changes in the outer membrane protein profile and the removal of iron from porcine transferrin. *Mol Microbiol* **3**, 1083-9.

Noinaj, N., Guillier, M., Barnard, T. J. & Buchanan, S. K. (2010). TonB-dependent transporters: regulation, structure, and function. *Annu Rev Microbiol* **64**, 43-60.

Nougayrède, J.-P., Taieb, F., De Rycke, J. & Oswald, E. (2005). Cyclomodulins: bacterial effectors that modulate the eukaryotic cell cycle. *Trends Microbiol* **13**, 103-10.

Nwugo, C. C., Gaddy, J. A., Zimbler, D. L. & Actis, L. A. (2010). Deciphering the iron response in *Acinetobacter baumannii*: A proteomics approach. *J Proteomics* **74**, 44-58. Elsevier B.V.

Odenbreit, S., Swoboda, K., Barwig, I., Ruhl, S., Bore, T., Koletzko, S. & Haas, R. (2009). Outer membrane protein expression profile in *Helicobacter pylori* clinical isolates. *Infect Immun* **77**, 3782-3790.

Odugbo, M., Odama, L., Umoh, J. & Lamorde, A. (2006). *Pasteurella multocida* pneumonic infection in sheep: Prevalence, clinical and pathological studies. *Small Ruminant Res* **66**, 273-277.

Ofek, I., Hasty, D. L. & Doyle, R. J. (2003). *Bacterial adhesion to animal cells and tissues.* US: ASM Press.

Ogunnariwo, J. A. & Schryvers, A. B. (2001). Characterization of a novel transferrin receptor in bovine strains of *Pasteurella multocida*. *J Bacteriol* **183**, 890-896.

Ohkawa, I., Shiga, S. & Kageyama, M. (1979). An esterase on the outer membrane of *Pseudomonas aeruginosa* for the hydrolysis of long chain acyl esters. *J Biochem* **86**, 643-56.

Ohnishi, T., Horiguchi, Y., Masuda, M., Sugimoto, N. & Matsuda, M. (1998). *Pasteurella multocida* toxin and *Bordetella bronchiseptica* dermonecrotizing toxin elicit similar effects on cultured cells by different mechanisms. *J Vet Med Sci* **60**, 301-305.

Okuyama, H., Itoh, H., Fukunaga, N. & Sasaki, S. (1984). Separation and properties of the inner and outer membranes of a psychrophilic bacterium, *Vibrio* sp. strain ABE-1. *Plant Cell Physiol* **25**, 1255-1264.

Oldfield, N. J., Donovan, E. a, Worrall, K. E., Wooldridge, K. G., Langford, P. R., Rycroft, A. N. & Ala'Aldeen, D. a a. (2008). Identification and characterization of novel antigenic vaccine candidates of *Actinobacillus pleuropneumoniae*. *Vaccine* **26**, 1942-54.

Olvera, A., Pina, S., Pérez-Simó, M., Oliveira, S. & Bensaid, A. (2010). Virulence-associated trimeric autotransporters of *Haemophilus parasuis* are antigenic proteins expressed in vivo. *Vet Res* **41**, 26.

Orme, R., Douglas, C. W. I., Rimmer, S. & Webb, M. (2006). Proteomic analysis of Escherichia coli biofilms reveals the overexpression of the outer membrane protein OmpA. *Proteins* 4269-4277.

Ou, Y.-Y., Gromiha, M. M., Chen, S.-A. & Suwa, M. (2008). TMBETADISC-RBF: Discrimination of beta-barrel membrane proteins using RBF networks and PSSM profiles. *Comput Biol Chem* **32**, 227-31.

O'Connor, C. D. & Hames, B. D. (2008). *Proteomics*. UK: Scion Publishing Ltd.

Pagès, J.-M., James, C. E. & Winterhalter, M. (2008). The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat Rev Microbiol* **6**, 893-903.

Pajón, R., Yero, D., Lage, A., Llanes, A. & Borroto, C. J. (2006). Computational identification of beta-barrel outer-membrane proteins in *Mycobacterium tuberculosis* predicted proteomes as putative vaccine candidates. *Tuberculosis* **86**, 290-302.

Papasotiriou, D. G., Markoutsa, S., Meyer, B., Papadioti, A., Karas, M. & Tsiotis, G. (2008). Comparison of the membrane subproteomes during growth of a new

Pseudomonas strain on lysogeny broth medium, glucose, and phenol. *J Proteome Res* **7**, 4278-88.

Patarakul, K., Lo, M. & Adler, B. (2010). Global transcriptomic response of *Leptospira interrogans* serovar Copenhageni upon exposure to serum. *BMC Microbiol* **10**, 31.

Paustian, M. L., May, B. J. & Kapur, V. (2001). *Pasteurella multocida* gene expression in response to iron limitation. *Infect Immun* **69**, 4109-4115.

Paustian, M. L., May, B. J. & Kapur, V. (2002). Transcriptional response of *Pasteurella multocida* to nutrient limitation. *J Bacteriol* **184**, 3734-3739.

Peak, I. R. A., Srikhanta, Y., Dieckelmann, M., Moxon, E. R. & Jennings, M. P. (2000). Identification and characterisation of a novel conserved outer membrane protein from *Neisseria meningitidis*. *FEMS Immunol Med Mic* **28**, 329-334.

Peng, X., Xu, C., Ren, H., Lin, X., Wu, L. & Wang, S. (2005). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of *Pseudomonas aeruginosa* responding to ampicilin, kanamycin, and tetracycline resistance. *J Proteome Res* **4**, 2257-65.

Perez, B. a, Planet, P. J., Kachlany, S. C., Tomich, M., Fine, D. H. & Figurski, D. H. (2006). Genetic analysis of the requirement for flp-2, tadV, and rcpB in Actinobacillus actinomycetemcomitans biofilm formation. *J Bacteriol* **188**, 6361-75.

Perkins-balding, D., Ratliff-griffin, M. & Stojiljkovic, I. (2004). Iron Transport Systems in *Neisseria meningitidis* . *Microbiol Mol Biol Rev* **68**, 154-171.

Petersen, S. K. & Foged, N. T. (1989). Cloning and expression of the *Pasteurella multocida* toxin gene , *toxA* , in *Escherichia coli*. *Infect Immun* **57**, 3907-3913.

Petersen, K. D., Christensen, H., Bisgaard, M. & Olsen, J. E. (2001). Genetic diversity of *Pasteurella multocida* fowl cholera isolates as demonstrated by ribotyping and 16S rRNA and partial *atpD* sequence comparisons. *Microbiol* **147**, 2739-2748.

Pevsner, J. (2009). *Bioinformatics and functional genomics*, 2nd edn. Canada: John Wiley & Sons Ltd.

Phadke, N. D., Molloy, M. P., Steinhoff, S. a, Ulintz, P. J., Andrews, P. C. & Maddock, J. R. (2001). Analysis of the outer membrane proteome of *Caulobacter crescentus* by two-dimensional electrophoresis and mass spectrometry. *Proteomics* **1**, 705-20.

Pham, T. K., Sierocinski, P., van der Oost, J. & Wright, P. C. (2010). Quantitative proteomic analysis of *Sulfolobus solfataricus* membrane proteins. *J Proteome Res* **9**, 1165-72.

Phillips, G. J. (2001). Green fluorescent protein - a bright idea for the study of bacterial protein localization. *FEMS Microbiol Lett* **204**, 9-18.

Pieper, R., Huang, S.-T., Robinson, J. M., Clark, D. J., Alami, H., Parmar, P. P., Perry, R. D., Fleischmann, R. D. & Peterson, S. N. (2009). Temperature and growth phase influence the outer-membrane proteome and the expression of a type VI secretion system in *Yersinia pestis*. *Microbiol* **155**, 498-512.

Pieper, R., Huang, S.-T., Parmar, P. P., Clark, D. J., Alami, H., Fleischmann, R. D., Perry, R. D. & Peterson, S. N. (2010). Proteomic analysis of iron acquisition, metabolic and regulatory responses of *Yersinia pestis* to iron starvation. *BMC Microbiol* **10**, 30.

Platz, G. J., Bublitz, D. C., Mena, P., Benach, J. L., Furie, M. B. & Thanassi, D. G. (2010). A tolC mutant of *Francisella tularensis* is hypercytotoxic compared to the wild type and elicits increased proinflammatory responses from host cells. *Infect Immun* **78**, 1022-31.

Poetsch, A. & Wolters, D. (2008). Bacterial membrane proteomics. *Proteomics* **8**, 4100-22.

Poobalane, S., Thompson, K. D., Diab, A., Ardó, L., Jeney, G. & Adams, A. (2008). Protein expression by *Aeromonas hydrophila* during growth in vitro and in vivo. *Microb Pathogenesis* **45**, 60-9.

Prado, M. E., Dabo, S. M. & Confer, a W. (2005). Immunogenicity of iron-regulated outer membrane proteins of *Pasteurella multocida* A:3 in cattle: molecular characterization of the immunodominant heme acquisition system receptor (HasR) protein. *Vet Microbiol* **105**, 269-80.

Prasadarao, N. V., Wass, C. a, Weiser, J. N., Stins, M. F., Huang, S. H. & Kim, K. S. (1996). Outer membrane protein A of *Escherichia coli* contributes to invasion of brain microvascular endothelial cells. *Infect Immun* **64**, 146-53.

Prasadarao, N. V., Blom, A. M., Villoutreix, B. O. & Linsangan, L. C. (2002). A novel interaction of outer membrane protein A with C4b binding protein mediates serum resistance of *Escherichia coli* K1. *J Immunol* **169**, 6352-60.

Proft, T. & Baker, E. N. (2009). Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease. *Cell Mol Life Sci* **66**, 613-35.

Rabilloud, T., Vaezzadeh, A. R., Potier, N., Lelong, C., Leize-Wagner, E. & Chevallet, M. (2009). Power and limitations of electrophoretic separations in proteomics strategies. *Mass Spectrom Rev* **28**, 816- 843.

Randall, A., Cheng, J., Sweredoski, M. & Baldi, P. (2008). TMBpro: secondary structure, beta-contact and tertiary structure prediction of transmembrane beta-barrel proteins. *Bioinformatics* **24**, 513-20.

Rasmussen, A. A., Eriksen, M., Gilany, K., Udesen, C., Franch, T., Petersen, C. & Valentin-Hansen, P. (2005). Regulation of *ompA* mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Mol Microbiol* **58**, 1421-9.

Ravaoarinoro, M., Ciurli, C., Toma, E. & Morisset, R. (1994). Rapid method for isolating detergent-insoluble outer membrane proteins from *Pseudomonas aeruginosa*. *Electrophoresis* **15**, 594-6.

Rebers, P. a, Jensen, a E. & Laird, G. a. (1988). Expression of pili and capsule by the avian strain P-1059 of *Pasteurella multocida*. *Avian Dis* **32**, 313-8.

Register, K. B., Lee, R. M. & Thomson, C. (1998). Two-color hybridization assay for simultaneous detection of *Bordetella bronchiseptica* and toxigenic *Pasteurella multocida* from swine. *J Clin Microbiol* **36**, 3342-6.

Remmert, M., Linke, D., Lupas, A. N. & Söding, J. (2009). HHomp--prediction and classification of outer membrane proteins. *Nucleic Acids Res* **37**, W446-51.

Repoila, F. & Darfeuille, F. (2009). Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol Cell* **101**, 117-31.

Rhoades, A. K. R. & Rimler, R. B. (1987). Capsular groups of *Pasteurella multocida* isolated from avian hosts. *Avian Dis* **31**, 895-898.

Rhomberg, T. a, Karlberg, O., Mini, T., Zimny-Arndt, U., Wickenberg, U., Röttgen, M., Jungblut, P. R., Jenö, P., Andersson, S. G. E. & Dehio, C. (2004). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of the bacterial pathogen *Bartonella henselae*. *Proteomics* **4**, 3021-33.

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**, 316-29.

Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H. & Zehfus, M. H. (1985). Hydrophobicity of amivo acid residues in globular proteins. *Science* **229**, 834-838.

Ross, R. F. (2007). *Pasteurella multocida* and its role in porcine pneumonia. *Anim Health Res Rev* **7**, 13-29.

Rossetti, C. a, Galindo, C. L., Lawhon, S. D., Garner, H. R. & Adams, L. G. (2009). *Brucella melitensis* global gene expression study provides novel information on growth phase-specific gene regulation with potential insights for understanding *Brucella*:host initial interactions. *BMC Microbiol* **9**, 81.

Rost, B. (2001). Review : protein secondary structure prediction continues to rise. *J Struct Biol*.

Rost, B., Yachdav, G. & Liu, J. (2004). The PredictProtein server. *Nucleic Acids Res* **32**, W321-6.

Rubingh, C. M., Bijlsma, S., Derks, E. P. P. a, Bobeldijk, I., Verheij, E. R., Kochhar, S. & Smilde, A. K. (2006). Assessing the performance of statistical validation tools for megavariate metabolomics data. *Metabolomics* **2**, 53-61.

Ruffolo, C. G. & Adler, B. (1996). Cloning, sequencing, expression, and protective capacity of the *oma87* gene encoding the *Pasteurella multocida* 87-kilodalton outer membrane antigen. *Infect Immun* **64**, 3161-7.

Ruffolo, C. G., Tennent, J. M., Michalski, W. P. & Adler, B. (1997). Identification, purification, and characterization of the type 4 fimbriae of *Pasteurella multocida*. *Infect Immun* **65**, 339-43.

Ruiz, N., Kahne, D. & Silhavy, T. J. (2006). Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol* **4**, 57-66.

Russell, W. & Herwald, H. (2005). *Concepts in bacterial virulence*. Switzerland: S. Karger AG.

Saier, M. H., Tran, C. V. & Barabote, R. D. (2006). TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res* **34**, D181-6.

Saier, M. H., Yen, M. R., Noto, K., Tamang, D. G. & Elkan, C. (2009). The Transporter Classification Database: recent advances. *Nucleic Acids Res* **37**, D274-8.

Sanchez, S., Mizan, S., Quist, C., Schroder, P., Juneau, M., Dawe, D., Ritchie, B. & Lee, M. D. (2004). Serological Response to *Pasteurella multocida* NanH Sialidase in Persistently Colonized Rabbits. *Clin Diagn Lab Immun* **11**, 825-834.

Sandal, I., Shao, J. Q., Annadata, S., Apicella, M. a, Boye, M., Jensen, T. K., Saunders, G. K. & Inzana, T. J. (2009). *Histophilus somni* biofilm formation in cardiopulmonary tissue of the bovine host following respiratory challenge. *Microbes Infect* **11**, 254-63.

Sandy, M. & Butler, A. (2009). Microbial iron acquisition: marine and terrestrial siderophores. *Chem Rev* **109**, 4580-95.

Shin, J.-H., Lee, H.-W., Kim, S.-M. & Kim, J. (2009). Proteomic analysis of *Acinetobacter baumannii* in biofilm and planktonic growth mode. *J Microbiol* **47**, 728-35.

Sauer, K. (2003). The genomics and proteomics of biofilm formation. *Genome Biol* **4**, 219.

Saxena, M. K., Singh, V. P., Kumar, A. A., Chaudhuri, P., Singh, V. P., Shivachandra, S. B., Biswas, A. & Sharma, B. (2006). REP-PCR analysis of *Pasteurella multocida* isolates from wild and domestic animals in India. *Vet Res Commun* **30**, 851-861.

Schulz, G. E. (2002). The structure of bacterial outer membrane proteins. *Biochim Biophys Acta* **1565**, 308-17.

Seydel, A., Gounon, P. & Pugsley, a P. (1999). Testing the "+2 rule" for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. *Mol Microbiol* **34**, 810-21.

Sharp, P. M. & Li, W.-hsiung. **(1987).** The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* **15**, 1281-1295.

Shayegh, J., Parviza, M. & Hejazi, M. S. (2010a). Diversity of caprine and ovine *Pasteurella multocida* isolates based on 16S rRNA gene sequencing. *Iran J Vet Res* **11**, 373-378.

Shayegh, J., Atashpaz, S. & Hejazi, M. S. (2008). Virulence gene profile and typing of ovine *Pasteurella multocida*. *Asian J Anim Vet Adv* **3**, 206-213.

Shayegh, J., Mikaili, P., Movassagh Gazani, M. H., Hossinzadeh, M. H., Maheri-sis, N., Beyzai, A. R. & Mikaili, N. (2010). Evaluation of biofilm formation in different capsular genotypes of *Pasteurella multocida* and the effects of *Matricaria camomilla* L. on it. *Middle-East J Sci Res* **5**, 5130-513.

Shen, H.-B. & Chou, K.-C. (2010). Gneg-mPLoc: a top-down strategy to enhance the quality of predicting subcellular localization of Gram-negative bacterial proteins. *J Theor Biol* **264**, 326-33.

Shivachandra, S. B., Kumar, a a, Amaranath, J., Joseph, S., Srivastava, S. K. & Chaudhuri, P. (2005). Cloning and characterization of *tbpA* gene encoding transferrin-binding protein (TbpA) from *Pasteurella multocida* serogroup B:2 (strain P52). *Vet Res Commun* **29**, 537-42.

Shivachandra, S. B., Kumar, a a, Gautam, R., Joseph, S., Saxena, M. K., Chaudhuri, P. & Srivastava, S. K. (2006a). Characterization of avian strains of *Pasteurella multocida* by restriction endonuclease and amplified fragment length polymorphism. *Res Vet Sci* **81**, 8-18.

Shivachandra, S. B., Kumar, a a, Gautam, R., Singh, V. P., Saxena, M. K. & Srivastava, S. K. (2006b). Identification of avian strains of *Pasteurella multocida* in India by conventional and PCR assays. *Vet J* **172**, 561-4.

Shivachandra, S. B., Kumar, a a & Chaudhuri, P. (2008). Molecular characterization of avian strains of *Pasteurella multocida* serogroup-A:1 based on amplification of repetitive regions by PCR. *Comp Immunol Microb* **31**, 47-62.

Sigrist, C. J. a, Cerutti, L., de Castro, E., Langendijk-Genevaux, P. S., Bulliard, V., Bairoch, A. & Hulo, N. (2010). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res* **38**, D161-6.

Skrabanek, L., Saini, H. K., Bader, G. D. & Enright, A. J. (2008). Computational prediction of protein – protein interactions. *Mol Biotechnol* **38**, 1-17.

Snijder, H. J. & Dijkstra, B. W. (2000). Bacterial phospholipase A: structure and function of an integral membrane phospholipase. *Biochim Biophys Acta* **1488**, 91-101.

Song, T., Mika, F., Lindmark, B., Liu, Z., Schild, S., Bishop, A., Zhu, J., Camilli, A., Johansson, J., & other authors. (2008). A new Vibrio cholerae sRNA modulates colonization and affects release of outer membrane vesicles. *Mol Biotechnol* **70**, 100-11.

St Michael, F., Vinogradov, E., Li, J. & Cox, A. D. (2005). Structural analysis of the lipopolysaccharide from *Pasteurella multocida* genome strain Pm70 and identification of the putative lipopolysaccharide glycosyltransferases. *Glycobiology* **15**, 323-33.

Steen, J. a J. a, Harrison, P., Seemann, T., Wilkie, I., Harper, M., Adler, B. & Boyce, J. D. (2010). Fis is essential for capsule production in *Pasteurella multocida* and regulates expression of other important virulence factors. *PLoS Pathog* **6**, e1000750.

Sthitmatee, N., Numee, S., Kawamoto, E., Sasaki, H., Yamashita, K., Takahashi, N., Kataoka, Y. & Sawada, T. (2008). Protection of chickens from fowl cholera by vaccination with recombinant adhesive protein of *Pasteurella multocida*. *Vaccine*.

Storz, G. & Hengge, R. (2011). *Bacterial stress response*, 2nd edn. US: ASM Press.

Stroud, D. A., Meisinger, C., Pfanner, N. & Wiedemann, N. (2010). Assembling the outer membrane. *Science* **328**, 831-832.

Stull, T. L., Mack, K., Haas, J. E., Smit, J. & Smith, a L. (1985). A comparison of techniques for isolation of the outer membrane proteins of *Haemophilus influenzae* type b. *Anal Biochem* **150**, 471-80.

Supek, F. & Vlahovicek, K. (2005). Comparison of codon usage measures and their applicability in prediction of microbial gene expressivity. *BMC Bioinformatics* **6**, 182.

Szafron, D., Lu, P., Greiner, R., Wishart, D. S., Poulin, B., Eisner, R., Lu, Z., Anvik, J., Macdonell, C., & other authors. **(2004).** Proteome Analyst: custom predictions with explanations in a web-based tool for high-throughput proteome annotations. *Nucleic Acids Res* **32**, W365-71.

Szöor, B., Ruberto, I., Burchmore, R. & Matthews, K. R. (2010). A novel phosphatase cascade regulates differentiation in *Trypanosoma brucei* via a glycosomal signaling pathway. *Genes Dev* **24**, 1306-16.

Söding, J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951-60.

Tabb, D. L., Vega-Montoto, L., Rudnick, P. A., Variyath, A. M., Ham, A. L., Bunk, D. M., Kilpatrick, L. E., Billheimer, D. D., Blackman, R. K., Cardasis, H. L., Carr, S. A., Clauser, K. R., Jaffe, J. D., Kowalski, K. A., Neubert, T. A., Regnier, F. E., Schilling, B., Tegeler, T. J., Wang, M., Wang, P., Whiteaker, J. R., Zimmerman, L. J., Fisher, S. J., Gibson, B. W., Kinsinger, C. R., Mesri, M., Rodriguez, H., Stein, S. E., Tempst, P., Paulovich, A. G., Liebler, D. C., Spiegelman, C. (2010). Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* **9**, 761-776.

Tamm, L. K., Arora, A. & Kleinschmidt, J. H. (2001). Structure and assembly of beta-barrel membrane proteins. *J Biol Chem* **276**, 32399-402.

Tamm, L. K., Hong, H. & Liang, B. (2004). Folding and assembly of beta-barrel membrane proteins. *Biochim Biophys Acta* **1666**, 250-63.

Tan, S., Tan, H. T. & Chung, M. C. M. (2008). Membrane proteins and membrane proteomics. *Proteomics* **8**, 3924-32.

Tang, X., Zhao, Z., Hu, J., Wu, B., Cai, X., He, Q. & Chen, H. (2009). Isolation, antimicrobial resistance, and virulence genes of *Pasteurella multocida* strains from swine in China. *J Clin Microbiol* **47**, 951-8.

Tatusov, R. L., Galperin, M. Y., Natale, D. a & Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* **28**, 33-6.

Taylor, P. D., Toseland, C. T., Attwood, T. K. & Flower, D. R. (2006). Beta barrel trans-membrane proteins: Enhanced prediction using a Bayesian approach. *Bioinformation* **1**, 231-233.

Thein, M., Sauer, G., Paramasivam, N., Grin, I. & Linke, D. (2010). Efficient subfractionation of Gram-negative bacteria for proteomics studies. *J Proteome Res* **9**, 6135-6147.

Tokuda, H. (2009). Biogenesis of outer membranes in Gram-negative bacteria. *Biosci Biotechnol Biochem* **73**, 465-473.

Tokuda, H. & Matsuyama, S.-I. (2004). Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta* **1693**, 5-13.

Tomich, M., Planet, P. J. & Figurski, D. H. (2007). The *tad* locus : postcards from the widespread colonization island. *Nat Rev Microbiol* **5**, 363-375.

Tommassen, J. (2010). Assembly of outer-membrane proteins in bacteria and mitochondria. *Microbiol* **156**, 2587-96.

Townsend, K. M., Dawkins, H. J. S. & Papadimitriou, J. M. (1997). REP-PCR analysis of *Pasteurella multocida* isolates that cause haemorrhagic septicaemia. *Res Vet Sci* **63**, 151-155.

Townsend, K. M., Frost, A. J., Lee, C. W., Papadimitriou, J. M. & Dawkins, H. J. S. (1998). Development of PCR assays for species- and type-specific identification of *Pasteurella multocida* isolates. *J Clin Microbiol* **36**, 1096-1100.

Tsirigos, K. D., Bagos, P. G. & Hamodrakas, S. J. (2011). OMPdb: a database of beta-barrel outer membrane proteins from Gram-negative bacteria. *Nucleic Acids Res* **39**, D324-31.

Tusnady, G. E., Doszta, Z. & Simon, I. (2005). PDB _ TM : selection and membrane localization of transmembrane proteins in the protein data bank. *Nucleic Acids Res* **33**, 275-278.

Tusnády, G. E., Kalmár, L. & Simon, I. (2007). TOPDB: topology data bank of transmembrane proteins. *Nucleic Acids Res* 1-6.

Tyler, C. D., Lichti, C. F., Diekman, A. B. & Foley, S. L. (2008). Evaluation of differentially expressed proteins following serum exposure in avian pathogenic *Escherichia coli*. *Avian Dis* **52**, 23-27.

Trémoulet, F., Duché, O., Namane, A., Martinie, B. & Labadie, J.-C. (2002). A proteomic study of *Escherichia coli* O157:H7 NCTC 12900 cultivated in biofilm or in planktonic growth mode. *FEMS Microbiol Lett* **215**, 7-14.

van Ulsen, P., Kuhn, K., Prinz, T., Legner, H., Schmid, P., Baumann, C. & Tommassen, J. (2009). Identification of proteins of *Neisseria meningitidis* induced under iron-limiting conditions using the isobaric tandem mass tag (TMT) labeling approach. *Proteomics* **9**, 1771-81.

Valavanis, I. K., Bagos, P. G. & Emiris, I. Z. (2006). Beta-barrel transmembrane proteins: Geometric modelling, detection of transmembrane region, and structural properties. *Comput Biol Chem* **30**, 416-24.

Veenstra, T. D. & Yates, J. R. (2006). *Proteomics for biological discovery*. USA: John Wiley & Sons Ltd.

Veith, P. D., Brien-simpson, N. M. O., Tan, Y., Djatmiko, D. C., Dashper, S. G. & Reynolds, E. C. (2009). Outer membrane proteome and antigens of *Tannerella forsythia*. *J Proteome Rese* **8**, 4279-4292.

Veken, W., Shah, N. H., Klaasen, P., Oudega, B. & Graaf, F. K. D. (1996). Binding of host iron-binding proteins and expression of iron-regulated membrane proteins by different serotypes of *Pasteurella multocida* causing haemorrhagic septicaemia. *Microb Pathogenesis* **21**, 59-64.

Vidakovics, M. L. P., Paba, J., Lamberti, Y., Andre, C., Sousa, M. V. D. & Rodriguez, M. E. (2007). Profiling the *Bordetella pertussis* proteome during iron starvation research. *J Proteome Res* **6**, 2518 - 2528.

Vinckx, T., Wei, Q., Matthijs, S., Noben, J.-P., Daniels, R. & Cornelis, P. (2011). A proteome analysis of the response of a *Pseudomonas aeruginosa* oxyR mutant to iron limitation. *Biometals* **24**, 523-32.

Viratyosin, W., Ingsriswang, S., Pacharawongsakda, E. & Palittapongarnpim, P. (2008). Genome-wide subcellular localization of putative outer membrane and extracellular proteins in *Leptospira interrogans* serovar Lai genome using bioinformatics approaches. *BMC Genomics* **9**, 1-15.

Virji, M., Makepeace, K., Ferguson, D. J., Achtman, M. & Moxon, E. R. (1993). Meningococcal Opa and Opc proteins: their role in colonization and invasion of human epithelial and endothelial cells. *Mol Microbiol* **10**, 499-510.

Voss, J. G. (1964). Lysozyme lysis of Gram-negative bacteria without production of spheroplasts. *J Gen Microbiol* **35**, 313-7.

Waldispuhl, J., Berger, B., Clote, P. & Steyaert, J.-M. (2006). Predicting transmembrane beta-barrels and interstrand residue interactions from sequence. *Proteins* **65**, 61-74.

Walker, S. L., Hill, J. E., Redman, J. A., Elimelech, M., Al, W. E. T. & Icrobiol, A. P. P. L. E. N. M. (2005). Influence of growth phase on adhesion kinetics of *Escherichia coli* D21g. *Appl Environ Microbiol* **71**, 3093-3099.

Walther, D. M., Rapaport, D. & Tommassen, J. (2009). Biogenesis of beta-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell Mol Life Sci* **66**, 2789-804.

Wandersman, C. & Delepelaire, P. (2004). Bacterial iron sources: from siderophores to hemophores. *Annu Rev Microbiol* **58**, 611-47.

Wang, H. C., Badger, J., Kearney, P. & Li, M. (2001). Analysis of codon usage patterns of bacterial genomes using the self-organizing map. *Mol Biol Evol* **18**, 792-800.

Wang, N., Mackenzie, L., De Souza, A. G., Zhong, H., Goss, G. & Li, L. (2007). Proteome profile of cytosolic component of zebrafish liver generated by LC-ESI

MS/MS combined with trypsin digestion and microwave-assisted acid hydrolysis. *J Proteome Res* **6**, 263-72.

Watson, P. J. & Davies, R. L. (2002). Outbreak of *Pasteurella multocida* septicaemia in neonatal lambs. *Vet Rec* **151**, 420-2.

Wedderkopp, A., Primal, S. V., Silva, S. & Little, P. B. (1993). Differences in protein expression of *Haemophilus somnus* grown under conditions of iron-restriction. *Vet Microbiol* **35**, 91-100.

Wheeler, D. L., Barrett, T., Benson, D. a, Bryant, S. H., Canese, K., Church, D. M., DiCuccio, M., Edgar, R., Federhen, S., & other authors. **(2005).** Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **33**, D39-45.

Weiner, J. H. & Li, L. (2008). Proteome of the *Escherichia coli* envelope and technological challenges in membrane proteome analysis. *Biochim Biophys Acta* **1778**, 1698-713.

Weiser, G. C., DeLong, W. J., Paz, J. L., Shafii, B., Price, W. J. & Ward, A. C. S. (2003). Characterization of *Pasteurella multocida* associated with pneumonia in bighorn sheep. *J Wildlife Dis* **39**, 536-544.

Westermeier, R. (2005). *Electrophoresis in practice*. Germany: Wiley-VCH.

Westermeier, R. & Naven, T. (2002). *Proteomics in practice: a laboratory manual of proteome analysis.* Germany: Wiley-VCH.

Wimley, W. C. (2002). Toward genomic identification of beta-barrel membrane proteins : Composition and architecture of known structures. *Protein Sci* **11**, 301-312.

Wu, Z., Feng, E., Wang, Y. & Chen, L. (2007). Discrimination of outer membrane proteins by a new measure of information discrepancy. *Protein Peptide Lett* **14**, 37-44.

Xiong, X.-P., Zhang, B.-W., Yang, M.-J., Ye, M.-Z., Peng, X.-X. & Li, H. (2010). Identification of vaccine candidates from differentially expressed outer membrane proteins of *Vibrio alginolyticus* in response to NaCl and iron limitation. *Fish Shellfish Immun* **29**, 810-6. Elsevier Ltd.

Xu, C., Wang, S., Ren, H., Lin, X., Wu, L. & Peng, X. (2005). Proteomic analysis on the expression of outer membrane proteins of *Vibrio alginolyticus* at different sodium concentrations. *Proteomics* **5**, 3142-52.

Yagupsky, P. & Slonim, A. (2005). Characterization and immunogenicity of *Kingella kingae* outer-membrane proteins. *FEMS Immunol Med Mic* **43**, 45-50.

Yan, C., Hu, J. & Wang, Y. (2008). Discrimination of outer membrane proteins with improved performance. *BMC Bioinformatics* **9**, 47.

Yu, C.-sheng & Lin, C.-jen. **(2004).** Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci* **13**, 1402-1406.

Yu, C.-sheng, Chen, Y.-ching, Lu, C.-hao & Hwang, J.-kang. **(2006).** Prediction of protein subcellular localization. *Proteins* **651**, 643- 651.

Yu, F., Anaya, C. & Lewis, J. P. (2007). Outer membrane proteome of *Prevotella intermedia* 17: identification of thioredoxin and iron-repressible hemin uptake loci. *Proteomics* **7**, 403-12.

Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S. C., Ester, M., & other authors. **(2010).** PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608-15.

Yu, N. Y., Laird, M. R., Spencer, C. & Brinkman, F. S. L. (2011). PSORTdb--an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea. *Nucleic Acids Res* **39**, D241-4.

Zhai, Y. & Saier, M. H. (2002). The beta-barrel finder (BBF) program , allowing identification of outer membrane beta-barrel proteins encoded within prokaryotic genomes. *Protein Sci* **11**, 2196-2207.

Zhou, X., Ren, L., Li, Y., Zhang, M., Yu, Y. & Yu, J. (2010). The next-generation sequencing technology: a technology review and future perspective. *Sci China Life Sci* **53**, 44-57.

Zou, H., Zheng, M., Luo, X., Zhu, W., Chen, K., Shen, J. & Jiang, H. (2008). Dynamic mechanism of fatty acid transport across cellular membranes through FadL: molecular dynamics simulations. *J Phys Chem B* **112**, 13070-8.

Zou, L. & Wang, Z. (2007). Predicting transmembrane topology of β-barrel membrane proteins with a hidden markov model. *IEEE* 145-148.